ALGEBRA

LECTURES DELIVERED TO POST-GRADUATE STUDENTS OF CALCUTTA UNIVERSITY

BY

FRIEDRICH WILHELM LEVI, DR.PHIL.NAT.
HARDINGE PROFESSOR

PART III—CONTINUED FRACTIONS

PART IV—APPROXIMATE SOLUTION

PART V—MATRICES RESULTANTS

PUBLISHED BY THE
UNIVERSITY OF CALCUTTA
1937

PRINTED IN INDIA

PRINTED AND PUBLISHED BY BHUPENDRALAL BANERJEE
AT THE CALCUTTA UNIVERSITY PRESS, SENATE HOUSE, CALCUTTA.

Reg. No. 1062B—December, 1937—E

PREFACE

The third and the fourth part of these lectures deal with some classical portions of Algebra. Obviously a strict selection has been necessary, and the author gave preference to those portions of Continued Fractions (Part III) which are connected with the theory of numbers. In Part IV (Approximation solution) much importance has been given to how to carry out the calculations. Of course this branch of Algebra has to be considered from quite a different point of view from that of General Algebra. By approximation numerical values should be found out in the quickest and easiest way; so the reader cannot get a real insight into the nature and importance of the different methods without knowing the difficulties a practical reckoner has to face. The hints given by the author for the simplification of calculations do not involve the use of slide rules, calculating machines and graphical methods as these expedients are not familiar to our students. In this as well in some other items a later edition of these lectures is expected to show some alteration.

Part V contains the most important theorems on matrices with application on Hermitian and quadratic forms. Here the student may have the satisfaction of seeing in a nutshell the greater part of what he learned on Analytic Geometry.

As in the prefaces of Part I and of Part II, I have much pleasure in delivering my thanks here to friends and kind collaborators. The Vice-Chancellor of our University, Syamaprasad Mookerjee, Esq., M.A., B.L., M.L.A., Barrister-at-Law, gave me every possible help to get these lectures printed in a very short time, and is fully entitled to the grateful thanks of the author as well of the students. Proofs and manuscripts have been carefully revised by Mr. A. C. Chowdhury, M.Sc., Research Scholar of Calcutta University. The Calcutta University Press kindly co-operated in carrying out the printing in the proposed time.

F. W. Levi

Calcutta, Asutosh Building,
October, 1937

INDEX

PART III. CONTINUED FRACTIONS

PART IV. APPROXIMATE SOLUTION

PART V. MATRICES. RESULTANTS

PART III

CONTINUED FRACTIONS

§ 1. General properties of continued fractions.

Let K be a field, S a subring of K, and A be a subset of K with the following property: If a class of residues $\neq (0)$ of S contains an element of A, this class contains also the inverse of an element of A. Hence if [1/1]

$$\alpha, \alpha', \alpha'', \dots, \alpha_1, \alpha_2, \dots \tag{1, 1}$$

denote elements of A and

$$s, s', s'', \dots, s_1, s_2, \dots \tag{1, 2}$$

denote elements of S, then every element of A can be represented either by

$$\alpha = s + 1 : \alpha' \tag{1, 3}$$

or by

$$\alpha = s. \tag{1, 3'}$$

If *e. g.*, K is the field of the real numbers, S the ring of the integers, and A the set of the real numbers ≥ 1, the representation (1,3), (1,3') is always possible and s is uniquely defined by α as the greatest integer $\leq \alpha$.

But the conditions (1,3), (1,3') hold also for other sets A in fields K, so an investigation in general terms is helpful.

Let

$$\begin{aligned} \alpha_1 &= s_1 + 1 : \alpha_2 \\ \alpha_2 &= s_2 + 1 : \alpha_3 \\ &\cdots\cdots\cdots\cdots \\ \alpha_n &= s_n + 1 : \alpha_{n+1}, \end{aligned} \tag{1, 4}$$

then

$$\alpha_1 = s_1 + \cfrac{1}{s_2 + \cfrac{1}{s_3 + \cfrac{}{\ddots \; s_n + \cfrac{1}{\alpha_{n+1}}}}}. \tag{1, 5}$$

The representation of α_1 by (5) is said to be a *continued fraction*. The formulae (1,4) can be continued till an element α_r will be an element of S. If there is an m so that $\alpha_m = s_m$, then the continued fraction is *finite*, otherwise it is *infinite*. If α can be represented by a finite continued fraction,

it belongs to the quotient field Q of S, and every finite set of elements

$$a_1, \dots, a_n \tag{1, 6}$$

of S defines an element of Q by the help of (1,5).

We now define other sequences of elements of S by the following formulas.

$$P_{-1}=0, \quad P_0=1, \quad P_k = a_k P_{k-1} + P_{k-2}, \quad k=1,2,\dots \tag{1, 7}$$

$$Q_{-1}=1, \quad Q_0=0, \quad Q_k = a_k Q_{k-1} + Q_{k-2}, \tag{1, 8}$$

$$D_k = \begin{vmatrix} P_k & P_{k-1} \\ Q_k & Q_{k-1} \end{vmatrix}, \tag{1, 9}$$

then from (1,7) (1,8) (1,9) it follows that

$$D_k = -D_{k-1}, \text{ and as } D_0 = 1, \ D_k = (-1)^k. \tag{1, 10}$$

From (1,9) and (1,10) it follows that P_k and Q_k have no other common factors in S than unities and that

$$P_k : Q_k - P_{k-1} : Q_{k-1} = (-1)^k : (Q_k Q_{k-1}). \tag{1, 11}$$

Let α_1 be an arbitrary element $\neq 0$ of K, then we get a uniquely defined sequence of elements $\alpha_1, \alpha_2, \dots, \alpha_{n+2}$ by the equations

$$a_i = \alpha_i : \alpha_{i+1} \tag{1, 12}$$

From (1,4) we get by multiplying the equations with $\alpha_2, \alpha_3, \dots$ respectively

$$\alpha_i = a_i \alpha_{i+1} + \alpha_{i+2}, \qquad i=1, \dots, n \tag{1, 4'}$$

From (1,4') (1,7) (1,8) we get

$$\begin{aligned} P_k \alpha_{k+1} + P_{k-1}\alpha_{k+2} &= P_{k-1}(a_k \alpha_{k+1} + \alpha_{k+2}) + P_{k-2}\alpha_{k+1} \\ &= P_{k-1}\alpha_k + P_{k-2}\alpha_{k+1}. \end{aligned}$$

By the repeated application of this formula we see that for $i<k$

$$P_k \alpha_{k+1} + P_{k-1}\alpha_{k+2} = P_i \alpha_{i+1} + P_{i-1}\alpha_{i+2} = P_0 \alpha_1 + P_{-1}\alpha_2 = \alpha_1, \tag{1, 13}$$

and by making a similar calculation with the elements Q we get

$$Q_k \alpha_{k+1} + Q_{k-1}\alpha_{k+2} = Q_i \alpha_{i+1} + Q_{i-1}\alpha_{i+2} = Q_0 \alpha_1 + Q_{-1}\alpha_2 = \alpha_2. \tag{1, 13'}$$

Hence

$$\begin{aligned} &(P_k a_{k+1} + P_{k-1}) : (Q_k a_{k+1} + Q_{k-1}) \\ &= (P_i a_{i+1} + P_{i-1}) : (Q_i a_{i+1} + Q_{i-1}) = a_1. \end{aligned} \tag{1, 13''}$$

If we multiply (1,13) with Q_{k-1} and (1,13') with $-P_{k-1}$ and add, then it follows from (1,10) that

$$(-1)^k \alpha_{k+1} = \alpha_1 Q_{k-1} - \alpha_2 P_{k-1}. \tag{1, 14}$$

From (1,12), (1,13'') and (1,10) it follows that

$$\alpha - \frac{P_k}{Q_k} = \frac{(-1)^{k-1}}{\alpha_2} \quad \frac{\alpha_{k+2}}{Q_k}. \tag{1, 15}$$

Hence if the continued fraction is finite, as $\alpha_{n+2} = 0$

$$\alpha_1 = \frac{P_n}{Q_n}. \tag{1, 15'}$$

If (1, 4) holds, the elements $e_1, e_2, \dots, e_n$ are said to be the *elements* of the continued fraction ; the quotients $P_k : Q_k$ are the *convergents* and α_{n+1} is called a *complete fraction*. As α_1 is uniquely defined by these elements, we shall denote α_1, if α_{n+1} exists, by

$$\alpha_1 = (e_1, \dots, e_n \mid \alpha_{n+1}) = \frac{P_n\alpha_{n+1} + P_{n-1}}{Q_n\alpha_{n+1} + Q_{n-1}}. \tag{1, 5'}$$

and if e_n is the last element of the continued fraction,

$$\alpha_1 = (e_1, \dots, e_n) = \frac{P_n}{Q_n}; \tag{1, 5''}$$

from (1,4), (1,5'), (1,5'') it follows that for $k \leqq n$

$$\alpha_k = (e_k, \dots, e_n \mid \alpha_{n+1}), \text{ respectively} \tag{1, $5_k'$}$$

$$\alpha_k = (e_k, \dots, e_n). \tag{1, $5_k''$}$$

Let P'_i and Q'_i be defined by

$$P'_{-1} = 0 \quad P'_0 = 1 \quad P'_i = e_{k+i-1} P_{i-1} + P_{i-2}$$

$$Q'_{-1} = 1 \quad Q'_0 = 0 \quad Q'_i = e_{k+i-1} Q_{i-1} + Q_{i-2}. \tag{1, 16}$$

then the following formulae hold :

$$\begin{vmatrix} P'_i & P'_{i-1} \\ Q'_i & Q'_{i-1} \end{vmatrix} = (-1)^i$$

$$\alpha_k = P'_i \alpha_{k+i} + P'_{i-1} \alpha_{k+i+1}$$

$$\alpha_{k+1} = Q'_i \alpha_{k+i} + Q'_{i-1} \alpha_{k+i+1}$$

$$(-1)^i \alpha_{k+i} = \alpha_k Q'_{i-1} - \alpha_{k+i-1} P'_{i-1} \tag{1, 17}$$

$$a_1 = \frac{P_n \text{[illegible]} + P_{n-1} \text{[illegible]}}{Q_n \text{[illegible]} + Q_{n-1} \text{[illegible]}}$$

From

$$P_n = a_n P_{n-1} + P_{n-2} \qquad Q_n = a_n Q_{n-1} + Q_{n-2}$$

$$P_{n-1} = a_{n-1} P_{n-2} + P_{n-3} \qquad Q_{n-1} = a_{n-1} Q_{n-2} + Q_{n-3}$$

$$P_1 = a_1 \qquad Q_2 = a_2$$

$$P_0 = 1 \qquad Q_1 = 1$$

we get the representation of $P_n : P_{n-1}$ and of $Q_n : Q_{n-1}$ as finite continued fractions. Then

$$\frac{P_n}{P_{n-1}} = (a_n, \ldots, a_2, a_1) \qquad \frac{Q_n}{Q_{n-1}} = (a_n, \ldots, a_2) \tag{1, 17a}$$

[1.2] There is a very close connection between the finite continued fractions and the algothimus of the *h.c.f.*

Let $a_1 : a_2$ be represented by a finite continued fraction $a_1 : a_2 = (a_1, \ldots, a_n)$. Then $a_n : a_{n+1} = a_n$. Hence $a_n = a_n a_{n+1} + 0$ therefore $a_{n+2} = 0$. From (1.11), (1.13), (1.13'), (1.14) we get therefore

$$a_1 = P_n a_{n+1}, \quad a_2 = Q_n a_{n+1}$$
$$(-1)^n a_{n+2} = a_1 Q_{n-1} - a_2 P_{n-1}. \tag{1, 18}$$

Hence $a_1 : a_2 = P_n : Q_n$ belongs to the quotientfield of S.

Every common factor of a_1 and a_2 is a factor of a_{n+1} and a_{n+1} is a common factor of a_1 and a_2. Hence a_1 and a_2 have an *h.c.f.* and this can be represented linearly by a_1 and a_2. Especially $a = P_n : Q_n$ is a representation by two relatively prime elements of S, as

$$P_n Q_{n-1} - Q_n P_{n-1} = (-1)^n.$$

Let every element of the quotientfield of S be representable by a finite continued fraction and let $a', a'' \neq 0$ be two arbitrary elements of S then $a' : a''$ can be represented by two relatively prime elements of S so that 1 can be represented in a linear and homogeneous manner by these elements.

Hence $a' : a'' = p : q$ and $pq' + qp' = 1$.

Therefore [illegible] and [illegible]

Hence $e' = $ [illegible]. So the arbitrary elements [illegible] of S have an h. c. f. [illegible] and this is represented in a [illegible] manner by e' and e''. From this consideration it follows that it is not possible to represent the quotients of the elements of an arbitrary [illegible] by finite continued fractions, e. g. if in S the elements are factorisable, but the factorisation is not unique there must be a quotient of two elements of S which can be represented by an infinite continued fraction only. As every element of S is represented by a finite continued fraction, the finite continued fractions do not form a field in these cases.

Let a function $N(e)$, which takes positive integral values only, be defined for every element $e \neq 0$ of S, and to every pair of elements e, e' of S let there exist two other elements e_1 and e'' so that

$$e = e_1 e' + e'' \qquad \text{and} \qquad \text{that}$$

$$\text{either} \qquad e'' = 0 \qquad \text{or} \qquad N(e'') < N(e'). \tag{1, 19}$$

Then [illegible] and as

$$N(e') > N(e'') > N(e''')\ldots > 0$$

are all integral numbers, the sequence $e', e'', \ldots$ must be finite. Hence [illegible] is a finite continued function. From these considerations we get the following theorem.

Theorem. Let S be an [illegible] containing 1 and let a positive integer $N(e)$ be defined satisfying the conditions (1, 19) for every element e of S, then the quotient field of S [illegible] formed by the finite continued fractions of S, and the highest common factor of two elements a_1 and a_2 of S is given by [illegible] in (1, 18), where [illegible] have the significance given by (1,6), (1,7), (1,8).

The formulae (1, 13) and (1, 14) are instances of [illegible] substitutions with coefficients from S, the determinant being ± 1.

Let A and B be the matrices of substitutions of this [illegible]

$$A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \qquad B = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} \qquad \begin{matrix} \det A = \varepsilon = \pm 1 \\ \det B = [illegible] = \pm 1 \end{matrix}$$

(1, 2[illegible])

$$A^{-1} = [illegible] \qquad B^{-1} = [illegible]$$

Let α be transformed by A into α', and α' be transformed by B into α'', then it follows

(1) α is transformed by E into α, and det E = 1

(2) α' is " " A' into α, and det $A' = \epsilon$

(3) α is " " AB into α'', and det AB $= \epsilon\epsilon' = \pm 1$

The product of two matrices has been defined in Part I, § 25 and it has been shown in § 28 that the determinant of a product of matrices is equal to the product of the determinants. The [illegible] which we have to consider here can easily be proved by direct calculation.

An element of K is said to be equivalent to α if we get it by transforming α by a linear fractional substitution with determinant ± 1. From (1), (2) and (3) it follows that this equivalence satisfies the conditions of reflexivity, symmetry and transitivity (see Part I, § 15) and therefore this equivalence defines a partition of K into classes, so that two elements of K are equivalent if and only if they belong to the same class.

By $\begin{pmatrix} s+1 & -1 \\ 1 & 0 \end{pmatrix}$ the element 1 becomes transformed into s; hence all elements of S are equivalent. From (1.13) and (1.5) it follows that the elements $x, x_1, x_2, \ldots$ defined by (1.4) are all equivalent. So in particular every finite continued fraction $x_1, \ldots, x_n$ is equivalent to $x_n = s_n$ and therefore belongs to the class containing the elements of S.

It is sometimes necessary to make a distinction between proper and improper equivalence. In the first case det A = 1, in the second case det A = −1. As det A = det A' in (1.20) both the notion of proper equivalence as well as the notion of improper equivalence is a symmetric one. By combining two equivalences of the same kind we get a proper equivalence and by combining two equivalences of different kind we get an improper equivalence. Every element is properly equivalent to itself, for the matrix E has the determinant 1. If in a class of equivalent elements an element α is also improperly equivalent to itself, i.e. if α is transformed into α by F, and det F = −1, then an arbitrary element β of the same class becomes transformed into α by H and by FH. One of these matrices has the determinant 1, the other has the determinant −1, so each element of the class is properly and improperly equivalent to α and therefore every element is at the same time properly and improperly equivalent to every other element of the class. If on the other hand α [illegible]

transformed into ξ by A as well as by B, where det A = 1, det B = −1, then ω becomes transformed into ω by BA, where det BA = −1, and therefore ω is improperly equivalent to itself, so that in this case every other element of the class C is properly and improperly equivalent to every other element of C. If in C there are no pairs of elements properly as well as improperly equivalent, then C must be divided in two classes without common elements: the elements of the 1st class are properly and the elements of the 2nd class are improperly equivalent to ω. Elements of the same subclasses are properly equivalent, elements of different subclasses are improperly equivalent.

Hence

Theorem. In a class of equivalent elements, either every element is properly and improperly equivalent to every other element or there are two subclasses without common element so that elements of different subclasses are improperly equivalent.

As 1 is transformed into itself by [illegible], every pair of elements is properly and improperly equivalent in the class containing the finite continued fractions.

Let ω be transformed into itself by A, then

$$\omega = \frac{a_1\omega + a_2}{a_3\omega + a_4} \text{ holds, hence } a_3\omega^2 + (a_4 - a_1)\omega - a_2 = 0. \tag{1, 21}$$

There are 3 different cases.

1. $a_3 = a_4 - a_1 = a_2 = 0$. In this case A = E or −E.

By these transformations every element is transformed into itself. E and −E generate proper equivalences.

2. (1, 21) is a reducible polynomial in ω. This is possible only if ω is an element of the quotientfield Q of B.

3. (1, 21) is irreducible. In this case ω is algebraic to Q and of order 2 over Q.

From these considerations it follows that not every element is improperly equivalent to itself.

Let the elements $\omega_1, \omega_2, \ldots$ as defined by (1, 1), be not all different. Let

$$\omega_\rho = \omega_\sigma \quad \text{then}$$

[illegible]

Hence [illegible] can be represented by an infinite periodic continued fraction with the period [illegible]. From (1, [illegible]) it follows that [illegible] becomes transformed into itself by the matrix $D = \begin{pmatrix} P'_t, & P'_{t-1} \\ Q'_t, & Q'_{t-1} \end{pmatrix}$ where det D [illegible] and becomes therefore equivalent to itself by this transformation. From (1, 13') it follows that $a_1 = \frac{P_{t-1}[illegible] + P_{t-2}}{Q_{t-1}a_t + Q_{t-2}}$ belongs to the field Q [illegible]. Hence [illegible] satisfies an equation of degree 2 with coefficients from S. The same holds for a_2, a_3 [illegible].

[1/4] Let now a periodic sequence [illegible]

of elements of S be given. It is not certain that in any extension of the quotient field Q of S there exists an element which can be represented by the infinite periodic continued fraction

$$(a_1, \ldots, a_m, a_1, \ldots, a_m, \ldots) \,. \qquad (1, 22)$$

If such an element exists, it is not certain that the element is uniquely defined in the field. But if there is a field in which there exists one and only one element represented by the periodic continued fraction (1, 22) then

$$a = (a_1, \ldots, a_m, a_1, \ldots, a_m, \ldots) = (a_1, \ldots, a_m \mid a)$$

holds and this is the case considered just before.

2. Representation of the positive numbers by continued fractions

[2/1] Let the elements [illegible] of the set [illegible] be the real numbers $\geqq 1$

and let S be the ring of the integers. Then the representation by the formulas (1, 3) and (1, 3')

$$a = s + 1 : a' \qquad \text{or} \qquad a = s$$

is always possible. If [illegible] is not an integral number, there is

$$1 \leqq s < a < s + 1, \quad a' = 1 : (a + 1 - s) [illegible]$$

and [illegible] representation [illegible] unique. [illegible] is an integral positive number, $n = a' + 1$, $a' = 0, 1, 2, \ldots$, there exist two possibilities

[illegible] (2, 1)

[illegible]

Therefore, [illegible] continued fraction

a_1 is a non-negative integral number, and

(2, 2)

$a_2, a_3, \ldots$ are positive integral numbers

[illegible] from [illegible] number [illegible] continued fraction. [illegible] integral number [illegible] there [illegible]

$$a_1 = (a' + 1)$$

and $$a_1 = a' + 1$$

If [illegible] integral [illegible] in the sequence $a_1, a_2, \ldots$ [illegible] integral and [illegible] fraction [illegible] first integral element [illegible] integral $1 \leq$ [illegible] continuation of the continued fraction :

$$a_{\nu+1} = (a' + 1), \text{ or } a_{\nu+1} = a', \; a_{\nu+2} = 1,$$

and there exist two and only two representations of a_1

$$a_1 = (a_1, \ldots, a_\nu, a' + 1), \text{ and } a_1 = (a_1, \ldots, a_\nu, a', 1).$$

If a_1 [illegible] representation [illegible]

If [illegible] ≤ 1 [illegible] > 1, and [illegible] considerations is given by the [illegible]

The [illegible] every positive [illegible] continued fraction satisfying the conditions [illegible] the

representation is unique and the continued fraction is infinite. If the number is rational, there exist two representations of it by an [illegible] finite continued fraction and there is no infinite continued fraction.

[2.2.] We are now going to prove the converse theorem, [illegible] every sequence satisfying [illegible] and [illegible]

Let $a_1, a_2, \ldots$ be an infinite sequence satisfying the conditions (1, 11). The numbers P and Q shall be defined by (1, 7) and (1, 8):

$$P_{-1} = 0,\ P_0 = 1,\ P_1 = a_1,\ P_2 = a_2 a_1 + 1;\ P_k = a_k P_{k-1} + P_{k-2}$$

$$Q_{-1} = 1,\ Q_0 = 0,\ Q_1 = 1,\ Q_2 = a_2;\ Q_k = a_k Q_{k-1} + Q_{k-2}$$

From P_k [illegible] it follows by mathematical induction that

$$Q_2, Q_3, \ldots, P_2, P_3, \ldots > 0 \quad \text{and that}$$

$$0 \leq P_1 < P_2 < P_3, \ldots < \ldots$$

$$Q_0 = 0 < Q_1 \leq Q_2 < Q_3, \ldots < \ldots \text{ hold} \tag{2, 3}$$

The quotients $\frac{P_k}{Q_k}$ are for $k \leq n$ the convergents of [illegible]

$(a_1, \ldots, a_k) = \frac{P_k}{Q_k}$. We can therefore apply (1, 15) on $x = \frac{P_k}{Q_k}$.

Hence

$$\frac{P_k}{Q_k} - \frac{P_{k+2}}{Q_{k+2}} = \frac{(-1)^{k+1} a_{k+2}}{Q_{k+2} \ Q_k} \quad \begin{matrix} > 0 & \text{if } k \geq 1 \text{ odd} \\ < 0 & ,,\ k > 1 \text{ even} \end{matrix}$$

Hence

$$\frac{P_1}{Q_1} < \frac{P_3}{Q_3} < \ldots < \frac{P_{2m-1}}{Q_{2m-1}}$$

$$\frac{P_2}{Q_2} > \frac{P_4}{Q_4} > \ldots > \frac{P_{2m}}{Q_{2m}} \qquad \text{for } m = 1, 2, 3, \ldots \tag{2, 4}$$

The quotients $\frac{P_k}{Q_k}$ form therefore two sequences, one is increasing, the other is decreasing and every number of the first sequence is less than every number of the second one. The intervals $\left(\frac{P_{2m-1}}{Q_{2m-1}}, \frac{P_{2m}}{Q_{2m}}\right)$ form therefore a

set of intervals, each of them is included in every other preceding it.

As $$\frac{P_n}{Q_n} - \frac{P_{n-1}}{Q_{n-1}} = \frac{(-1)^n}{Q_n Q_{n-1}}, \quad \text{[see (1, 3)]}$$

The length of the intervals converges to 0. Hence there is for every given sequence (a_i) one and only one real number α such that

$$\frac{P_{2m}}{Q_{2m}} < \alpha < \frac{P_{2m+1}}{Q_{2m+1}} \quad \text{holds for every } m \geq 1. \qquad (2, 5)$$

α is a positive number. This is the number which corresponds uniquely to the sequence $a_1, a_2, \dots$ and which is represented by the continued fraction. We will prove that the convergents of such fraction are equal to $a_1, a_2, \dots$. It is not obvious from the above text; only this is true, that if there exists a number which is represented as a continued fraction $(a_1, a_2, \dots)$, then this number can only be the number defined by that sequence, as there is no other number

situated within all the intervals $\left(\frac{P_n}{Q_n}, \frac{P_{n+1}}{Q_{n+1}}\right)$, but it may so happen that

there is no such number and that the representation of α as a continued fraction furnishes another sequence which defines the same real number α. Of course this complication cannot occur. We now show the following lemma which gives us some idea of the distribution of the continued fractions on the axis of the real numbers.

Lemma. Let $P_n : Q_n$ be the convergents of $(a_1, a_2, \dots)$, and $P' : Q'$ be the last convergent of $a_1, \dots, a_{n-1}, a_n + t$, where $n = 1, 2, \dots$, and $t > 0$, then

$$P_n : Q_n < P_{n+1} : Q_{n+1} \leq P' : Q' \qquad \text{if } n \text{ is odd,}$$
$$(2, 6)$$
and $$P_n : Q_n > P_{n+1} : Q_{n+1} \geq P' : Q' \qquad \text{if } n \text{ is even.}$$

Proof. $$Q_{n+1} = a_{n+1} Q_n + Q_{n-1}$$

$$Q' = (a_n + t)\, Q_{n-1} + Q_{n-2} = Q_n + t\, Q_{n-1}$$

$$\frac{1}{t} Q' = \frac{1}{t} Q_n + Q_{n-1} \leq Q_{n+1}$$

†The equality holds only if $n = 0$ or if $t = a_{n+1} = 1$.

Applying (1, 11) we get

$$[illegible] \quad (2, 7)$$

$$[illegible]$$

$$[illegible] \quad (2, 7)$$

From [illegible]

[illegible] have a common end point [illegible] However [illegible] number which is defined as an infinite continued fraction [illegible] every [illegible] sequence satisfying (1, 1) [illegible] theorem.

[illegible] (1, 1) is furnished by the representation of a [illegible] different real numbers correspond to different sequences [illegible] defined by the inequalities (2, 5).

[2.3] From the [illegible] the continued fractions are distributed on the axis of the real numbers [illegible] we will write for an arbitrary finite set $a_1 \ldots a_n$

$$(\ldots) = [n \mid t] \qquad t = 0, 1, 2 \ldots$$

$$(\ldots) = [n, s] \qquad s = 1, 2, \ldots$$

Let n be odd

$$[n, 1] < [n, 3] < \ldots < [n, 2] < [n-1] = [n \mid 1] < [n \mid 2] < \ldots$$

If n is even, the relation $<$ must be replaced by $>$. For the continued fraction having only one element [illegible], then [illegible]. So we get a first partition of the [illegible] by the [illegible]. The continued fractions [illegible] beginning with s_1 are [illegible] and s_1+1.

For the continued fraction [illegible] there [illegible] having [illegible] limiting point. The continued fraction [illegible] is situated between $3+\frac{1}{2}$ and [illegible] $3+$ [illegible] the limiting point is the [illegible]. [illegible] the intervals become [illegible] every segment [illegible] of the [illegible] divided by the [illegible] $n+1$ [illegible], [illegible] in every segment only one limiting point and [illegible].

$P_1=$ [illegible] P_2 [illegible] Q [illegible]

$Q_1=1 \quad Q_2=2 \quad Q_3=2s_3+1$

$s_3=$	1	2	3	4
[illegible] $s_3+1)=$ [illegible]	[illegible]	[illegible]	[illegible]	[illegible]

As every rational number is represented by [illegible] continued fraction, and [illegible] by a continued fraction with an [illegible] with an odd number of elements [illegible] the limiting point of one odd and one even [illegible]. So we get a natural classification of the rational [illegible] numbers. Numbers being represented [illegible] P_1, Q_1 [illegible] being represented by P_2, Q_2, etc.

The importance of this [illegible] the following theorem.

Theorem. If $s>0$, and $\frac{P_{2n-1}}{Q_{2n-1}} < \frac{r}{s} < \frac{P_{2n}}{Q_{2n}}$ then $s > Q_{2n} > Q_{2n-1}$.

Proof. From the supposition it follows directly that

$$0 < \frac{r}{s} - \frac{P_{2n-1}}{Q_{2n-1}} < \frac{P_{2n}}{Q_{2n}} - \frac{P_{2n-1}}{Q_{2n-1}} = \frac{1}{Q_{2n}Q_{2n-1}}$$

and as s and Q_{2n-1} are positive

$$0 < rQ_{2n-1} - sP_{2n-1} < \frac{s}{Q_{2n}}$$

the middle part of this inequality is an integral positive number.

Hence $1 < \frac{s}{Q_{2n}}$ i.e. $Q_{2n} < s$.

[illegible]

[illegible] by the [illegible] are the [illegible]

§ 3. PERIODIC CONTINUED FRACTIONS WITH INTEGRAL COEFFICIENTS.

3/1) [illegible] continued fraction in S [illegible] quadratic [illegible] S is the ring of the [illegible] that the periodic continued fractions [illegible]

Theorem. If a [illegible] belongs to a field [illegible] where [illegible] is represented by a periodic continued fraction.

Proof. α has to be the root of polynomial

$$a\alpha^2 + 2b\alpha + c \, ; \qquad (3, 1)$$

[illegible] the other hand [illegible] represented by a continued fraction $\alpha = (a_1, \ldots, a_n, \lambda)$. From (1, 13') it follows that

[illegible]

[illegible] $a(P_n\lambda + P_{n-1})^2 +$ [illegible] $(Q_n\lambda + Q_{n-1})^2$ [illegible]

or λ is a root of [illegible]

where [illegible]

Let β be the second root of [illegible] and μ be defined by

[illegible] hence [illegible]

$$\beta = \frac{P_n\mu + P_{n-1}}{Q_n\mu + Q_{n-1}} .$$

As [illegible] it follows that

$$\mu = \text{[illegible]}$$

As [illegible] $\frac{P_n}{Q_n}$ [illegible] $\frac{P_{n+1}}{Q_{n+1}}$ [illegible] $\frac{P}{Q_n} = \frac{1}{Q_n Q_{n+1}}$ [illegible] $\frac{1}{Q_n^2}$

$$\frac{P_n}{Q_n} + \frac{\epsilon}{Q_n^2}, \text{ where } |\epsilon| < 1.$$

So $$P_n - Q_n\alpha = \frac{\epsilon}{Q_n}.$$

From (3, 2) it follows therefore

[illegible]

$$< 0 \quad \text{for } Q_n Q_{n-1} |\beta - \alpha| > 2$$

Hence after a certain integer [illegible] negative [illegible] therefore [illegible] becomes negative. As [illegible] admits only a finite number of solutions [illegible] there must be different complete fractions satisfying the same equation [illegible] with [illegible]. These equations have only one positive root. Therefore the same complete fraction will be repeated. Hence the continued fraction is periodic.

The period of a continued fraction will be denoted by a bar, e.g. [4/21]

$$(a_1, \ldots, a_m, \overline{a_{m+1}, \ldots}) \qquad (3, 4)$$

denotes a continued fraction with the period [illegible]. If [illegible] the continued fraction is said to be purely periodic.

Let [illegible] be the complete fractions of [illegible]. As [illegible] for [illegible] every property which holds for complete fractions of sufficient high index holds for every complete fraction of a purely periodic continued fraction. From [illegible] we know that the conjugate of a complete fraction of sufficiently high index is negative. Therefore this property holds for every complete fraction of index [illegible] for every purely periodic continued fraction. In a purely periodic continued fraction [illegible] hence the continued fraction represents a number > 1.

A root [illegible] of a quadratic equation is called reduced if [illegible] > 1 and the conjugate root satisfies $0 > \alpha' > -1$. We will prove now that every purely periodic continued fraction represents a reduced quadratic number.

Let

$$\alpha = (a_1, \ldots, a_n) = (a_1, \ldots, a_n \mid \alpha) \qquad (3, 7)$$

$$\xi = (a_n, \ldots, a_1) = (a_n, \ldots, a_1 \mid \xi) \qquad (3, 7')$$

be two purely con[illegible] f[illegible]ons, the elements a_i being the same in both continued fractions but ordered in an inverse manner.

Let P_n, Q_n be the convergents of α. Then (see 1, 17°)

[illegible] $= (a_n, \ldots, a_1)$

hence [illegible]

Let $\beta =$ [illegible], then $0 > \beta > -1$, and

[illegible] hence [illegible] a root of

$$f(x) = Q_n x^2 + (Q_{n-1} - P_n) x - P_{n-1}$$

[illegible] also a root $f(x)$, and as $\alpha > 1$,

[illegible] consider [illegible] therefore

[illegible] fraction (3, 7) [illegible] conjugate to α, and $\xi = -\beta^{-1}$, then ξ is represented by (3, 7')

Every continued fraction [illegible] complete [illegible] equivalent to a purely continued fraction; hence

Corollary. Every quadratic number is equivalent to a reduced quadratic number

[illegible]

$$\alpha = \frac{a + \sqrt{D}}{b} = s + [illegible], \text{ where } s < \alpha < s + 1,$$

and $a, b, s, D > 0$ are integral numbers.

[illegible] $b' = (D - a'^2) : b$

From these formulas we can find out by a simple procedure the numbers a, b, c defining uniquely the complete fractions [illegible] and the numbers [illegible] defining the continued fraction. As the continued fraction is periodic, one period must be repeated after a finite number of steps. Then the first period is finished and the calculation has to be stopped.

Examples.

1. $\alpha = \dfrac{1+\sqrt{5}}{2}$ (harmonic section) $D=5$

a	b		c
1	2	$0 < \dfrac{1+\sqrt{5}}{2} < 1$ [illegible]	1
1	[illegible]	$1 < \dfrac{1+\sqrt{5}}{2} < 2$	1
1	2		

The last complete fraction is therefore equal to the preceding; hence $\alpha = (1, \dot{1})$. This is the simplest continued fraction, but the worst for practical calculation [illegible] the numbers P_n, Q_n are increasing slower than in any other case.

2. $\alpha = \sqrt{26}$.

a	b	c
0	1	5
	1	10
5	1	

hence $\alpha = (5, \dot{10})$. This example is very convenient for quick and exact calculation.

$$
\begin{array}{llll}
P_0 = & 1 & Q_0 = & 0 \\
P_1 = & 5 & Q_1 = & 1 \\
P_2 = & [illegible] & Q_2 = & 10 \\
P_3 = & [illegible] & Q_3 = & 101 \\
P_4 = & [illegible] & Q_4 = & 1020 \\
P_5 = & [illegible] & Q_5 = & 10301 \\
P_6 = & [illegible] & Q_6 = & 104030
\end{array}
$$

A

Hence $\alpha = \frac{[illegible]}{[illegible]} - \epsilon$, where $0 < \epsilon < 10^{-13}$. Therefore $\alpha = [illegible]$, the last two figures being uncertain.

As the error $\epsilon = \left|\alpha - \frac{P_n}{Q_n}\right| < \frac{1}{Q_n Q_{n+1}} < \frac{1}{a_{n+1} Q_n^2}$, it is useful to stop the calculation just before a large a_n.

([illegible]) $\alpha = \sqrt{[illegible]}$, $D = 2$

a	b	c
0	1	1
1	1	[illegible]
1	1	

Hence $\sqrt{2} = [illegible]$. As P_n, Q_n are increasing very slowly we will use another method.

$$\sqrt{2} = \frac{\sqrt{200}}{10}. \quad \text{If } \sqrt{200} = \frac{P_n}{Q_n} \pm \epsilon, \quad \sqrt{2} = \frac{P_n}{10Q_n} \pm \frac{\epsilon}{10}$$

So we represent $\sqrt{200}$ by a continued fraction. $D = 200 = 14^2 + 4$.

a	b		c
0	1	$14 < \sqrt{200} < 15$	14
14	4	$7 < \frac{14 + \sqrt{200}}{4} < 8$	7
14	1	$28 < 14 + \sqrt{200} < 29$	28
14	4		

Hence $\sqrt{200} = (14, 7, 28)$.

$$\begin{array}{llll} P_0 = & 1 & Q_0 = & 0 \\ P_1 = & 14 & Q_1 = & 1 \\ P_2 = & 99 & Q_2 = & 7 \\ P_3 = & 2786 & Q_3 = & 197 \\ P_4 = & 19601 & Q_4 = & 1386 \end{array}$$

$$\sqrt{200} = \frac{19601}{1386} - \epsilon, \quad 0 < \epsilon < \frac{1}{Q_4 Q_5} < \frac{1}{28\, Q_4^2} < 3 \cdot 10^{-8}$$

$$\sqrt{2} = \frac{19601}{13860} - \epsilon', \quad 0 < \epsilon' < 3 \cdot 10^{-9} = 1.41421356[illegible]$$

true to eight figures after the decimal point.

Exercise. Prove that [illegible] $= \sqrt{a^2+1}$, and calculate $\sqrt{2+1}$, $\sqrt{82}$, $\frac{\sqrt{i+2}}{8}$, $\sqrt{17}$. Calculate $\sqrt{3}$ directly and also by help of $\sqrt{}$ [illegible].

In order to prove the converse of the last theorem, it is useful to consider the following lemma.

Lemma. If $\alpha > 1$ and $-1 < \beta < 0$ are conjugate quadratic numbers, $\alpha = (a, a_1, a_2, \ldots)$, then all the complete fractions $\alpha_1 = (a_1, a_2, \ldots)$, $\alpha_2 = (a_2, \ldots)$, are reduced numbers.

Proof. $\alpha = a + \frac{1}{\alpha_1}$, $\beta = a + \frac{1}{\beta_1}$, α_1 and β_1 are conjugate numbers, $\alpha_1 > 0$, $-\frac{1}{\beta_1} = a - \beta > a \geq 1$ hence α_1 is reduced, and by repetition of this procedure it follows that $\alpha_2, \ldots$ are reduced.

Theorem. Every reduced quadratic number is represented by a purely periodic continued fraction.

Proof. Every quadratic number is represented by a periodic continued fraction $\alpha = ($ [illegible] $a, a_1, \ldots a_n)$. Let this number be reduced and let the periodicity of the continued fraction begin with [illegible] only [illegible] let $a \neq a_n$, then it follows from the last lemma that $\overline{a, a_1, \ldots}$ is a reduced number [illegible]. We will prove that this is impossible. Using the same notations as in the lemma we state

$\alpha_1 = \alpha_{n+1}$, hence $\beta_1 = \beta_{n+1}$

$\alpha = a + \frac{1}{\alpha_1}$, $\alpha_n = a_n + \frac{1}{\alpha_{n+1}}$, hence $\beta = a + \frac{1}{\beta_1}$, $\beta_n = a_n + \frac{1}{\beta_{n+1}}$

$$-\frac{1}{\beta_1} = a - \beta, \quad -\frac{1}{\beta_{n+1}} = a_n - \beta_n,$$

but as α and α_n are reduced $0 < -\beta < 1$ and $0 < -\beta_n < 1$ hold, hence $a < -\frac{1}{\beta_1} < a+1$ and $a_n - 1 < \frac{-1}{\beta_{n+1}} < a_n$. From $\beta_1 = \beta_{n+1}$ it follows therefore that $a = a_n$.

Theorem. Let $\alpha = \sqrt{\frac{p}{q}} > 1$ be rational, then [3/5]

$$\alpha = (a, a_1, \ldots a_n, \ldots) \qquad (5, 8)$$

and $s_n = 2s$ [illegible]

$$s_i = s_{n-i} \qquad i = 1, \ldots, n-1 \tag{8, 8'}$$

hold. [illegible] and [illegible] hold, then [illegible] irrational square root > 1.

Proof. As [illegible] > 1 [illegible] the number [illegible] < 0 [illegible] conjugate to [illegible] it follows from the lemma that [illegible] from [illegible] s_1, s_2 [illegible] are reduced and therefore [illegible]. Hence [illegible] satisfies [illegible]. As [illegible]

and [illegible] and [illegible] are conjugate

$$\alpha_1 = (s_1, \ldots, s_n) \tag{8, 9}$$

Therefore it follows from the lemma [illegible], that

$$-1 : \beta_1 = (s_n, \ldots, s_1) \tag{8, 9'}$$

$0 =$ [illegible] $+ 1$ [illegible]. Hence [illegible] $- 1$ [illegible]

Therefore $\quad ($ [illegible] $) = (2s,$ [illegible] $) \tag{8, 10}$

holds. [illegible] is equivalent to [illegible] (8, 10) [illegible] defined by (8, 8), (8, 8'), then (8, 10) holds.

Let s_1 and [illegible] be defined by [illegible] and let $\beta =$ [illegible] 1 [illegible], then it follows from [illegible] and [illegible] that [illegible] and [illegible] are [illegible] and therefore [illegible] are conjugate [illegible] from [illegible] it follows that [illegible] $= 0$. Hence α and β are the roots of a rational [illegible] equation [illegible]. From [illegible] $= 2s$ it follows that [illegible] ≥ 1, and therefore [illegible] > 1 and as [illegible] (8, 9) is an infinite continued fraction, [illegible] irrational.

Corollary. Let [illegible] and P_k [illegible] Q_k [illegible] be the convergents of [illegible], then

$$t P_{kn}^2 - r Q_{kn}^2 = (-1)^{kn} t \tag{8, 11}$$

holds for every $k = 1, 2, \ldots$

Proof. Let [illegible] be the complete fractions of [illegible], then [illegible]

$$[illegible] = [illegible] + \frac{1}{[illegible]} = 2s + \frac{1}{[illegible]}, \ldots$$

But as $\omega = \dfrac{P_{k-1} \ldots + P_{k-1}}{Q_{k-1} \ldots + Q_{k-1}}$, [illegible] 13

$$\omega = \frac{P_{kn}(\ldots + \omega) + P_{k\ldots}}{Q_{kn}(\ldots + \omega) + Q_{k\ldots}}$$ [illegible], hence

$$Q_{kn}\,\omega^2 - P_{kn}\,\ldots - P_{kn-1} + \omega(Q_{kn}\,\ldots + Q_{\ldots}\ \ldots) = 0$$

As [illegible] $= \ldots : 1$ is rational, and ω is irrational

$$P_{k\ldots 1} + P_{k\ldots}\ \ Q_{k\ldots}\ [illegible]$$

$$= P_{kn} + Q_{kn}\,\ldots + Q_{kn-1}\ \ldots$$

hold. If we multiply these equations by $Q_{\ldots}$, [illegible] respectively $-P_{kn}$, and add, we get $(P^2_{kn} - \ldots Q^2_{kn} + \ldots(P_{k\ldots 1}\,Q_{kn} - Q_{k\ldots 1}\,P_{kn}) = 0$ and from this formula we get (3, 11) directly.

§ 4. Applications on Theory of Numbers

It is proposed to solve [4, 1]

$$[illegible] \qquad (4, 1)$$

by integral x and y.

Of [illegible] (4, 1) cannot be solved if there is a common factor of a and [illegible] different from [illegible]. We therefore [illegible] a and b [illegible] to be relatively prime. a [illegible] b can be represented by an even [illegible] continued fraction as [illegible]

a [illegible] $b = (\ldots_1, \ldots, \ldots_{2n})$

a [illegible] $b = P_{2n}$ [illegible] Q_{2n}, and as a and b are positive and relatively prime

$a = P_{2n}$, $b = Q_{2n}$ and therefore

$$a\,Q_{2n-1} - b\,P_{2n-1} = P_{2n}\,Q_{2n-1} - Q_{2n}\,P_{2n-1} = (-1)^{2n} = 1$$

holds. Hence we get the integral solutions by

$$x = Q_{2n-1} + k\,b,$$
$$y = P_{2n-1} + k\,a \qquad \text{where } k = 0, \pm 1, \pm 2, \ldots$$

To solve $x^2 - dy^2 = 1$ (Pell's equation) by integral x and y, [illegible] and use (4, 2) (3, 11).

$$\sqrt{d} = \omega = \ldots, \ldots_1 \ldots \ldots_n$$

then $\quad P^2_{kn} - d\,Q^2_{kn} = (-1)^{kn}$

107,999

Therefore if n is even, $\quad (x, y) = (P_{kn-1}, Q_{kn-1})$

and if n is odd, $\quad (x, y) = (P_{2kn-1}, Q_{2kn-1})$

are solutions for every positive integer k.

E.g. $\quad x^2 - 26\, y^2 = 1$

$$\sqrt{26} = (5, \dot{10}) \qquad n = 1$$

By this method we get the solutions

$$(x, y) = (P_1, Q_1) = (51, 10)$$
$$= (P_3, Q_3) = (5201, 1020)$$
$$= (P_5, Q_5) = (530451, 104030)$$

§ 5. CONSTRUCTION OF THE SYSTEM WHOSE ELEMENTS ARE $\phi(x)$

[5, 1] In § 1 and 2 the system of the positive numbers has been taken for the system A, S being the ring of the integers. We will now consider another system A.

Let K be an arbitrary field and x an indeterminate not included in K. The elements of K will be denoted by $a, b, \ldots, c$ with and without indices. The elements of the ring K[x] will be denoted by

$$f(x),\ g(x), \tag{5, 1}$$

The ring (a ring of a field) K[x] will be set as the ring S.

In order to get a new system A we create new elements denoted by Greek letters

$$\phi(x),\ \psi(x),\ \chi(x),\ \omega(x) \tag{5, 2}$$

in the following manner.

$$\phi(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0 + a_{-1} x^{-1} + \ldots + a_{-k} x^{-k} + \ldots + [illegible]$$

$$= [illegible] = \sum_{i=-\infty}^{n} a_i x^i \tag{5, 3}$$

This is a purely formal definition. It means that to every sequence of coefficients from K with fixed decreasing integral indices

then $\frac{\phi}{\psi}$ is equal to one of our new elements and this element will not be changed if we take before a_n a finite set of nullcoefficients. Nothing has been supposed about convergence. We have to define the addition and the multiplication of the elements $\phi(x)$ and we will define them in such a way that the elements $\phi(x)$ for which $a_k = 0$ for $k < 0$ form a subring isomorphic to K $[x]$.

So we define :

Let $n \geq m$, $\phi(x) = \sum_{k=n}^{-\infty} a_k x^k$

$$\psi(x) = \sum_{k=m}^{-\infty} b_k x^k = \sum_{k=n}^{-\infty} b_k x^k \quad \text{where } b_n = \dots = b_{m+1} = 0 \text{ for } n > m$$

then $\phi(x) + \psi(x) = \chi(x) = \sum_{k=n}^{-\infty} c_k x^k$

$$\phi(x) \cdot \psi(x) = \omega(x) = \sum_{k=n+m}^{-\infty} d_k x^k$$

where $\qquad c_k = a_k + b_k, \qquad d_k = \Sigma a_i \, b_{k-i} \qquad (5, 4)$

$$n \geq i \geq k - m.$$

The definitions (5, 4) are obviously independent of nullcoefficients put before; the commutative, associative and distributive laws hold, and the subtraction is uniquely defined by

$$b_k = c_k - a_k$$

Hence for the nullelement every coefficient is 0. If for the coefficients of $\psi(x)$ the conditions $b_k = 0$, for $k \neq 0$ hold, then in $\phi(x)\,\psi(x) = \sum d_k x^k$

$$d_k = b_0 \, a_k \qquad \text{holds.}$$

The elements $\phi(x)$ form a ring R and those elements for which $a_k = 0$, $k < 0$ form a subring for which the addition and multiplication has been defined in the same way as for polynomials. Hence there is an isomorphism I by which this subring becomes isomorphic to K $[x]$. Let $\phi(x) \neq 0$, then $\phi(x)$ has at least one coefficient $\neq 0$, let n be the highest index of the non vanishing co-efficients, then

$$\phi(x) = \sum_{k=n}^{-\infty} a_k x^k, \; a_n \neq 0$$

[illegible] the degree of [illegible]. From [illegible] follows directly

The degree of a product [illegible] equal to the sum of the degrees of the factors.

The degree of a sum of elements [illegible] different degrees is equal to the maximum degree of the summands.

[illegible] a product [illegible] two elements [illegible] cannot [illegible]. Hence the ring [illegible]. We will now [illegible] the elements [illegible] with the corresponding elements [illegible]

So the elements [illegible] identified with b_k

and $b_k\,\varphi(x) = \sum b_k a_k x^k$ holds.

[illegible] then [illegible]. Every field [illegible] the elements of [illegible] which are [illegible] form a ring which [illegible] to [illegible] ring of Q. [illegible] elements of this ring will therefore be identified with the corresponding elements of Q. [illegible] becomes identified with [illegible] and the [illegible] sum [illegible] becomes identified with the [illegible] sum $\varphi(x) = \sum [illegible]\, x^k$, where $a_k = 0$, for $k < -m$.

[illegible] we can extend the [illegible] of division of the polynomials to the elements (5, 2)

[illegible] and [illegible] of degree n and m respectively and [illegible] and b_m their highest co-efficients.

[illegible] of degree $n_1 < n$

By repetition of this procedure we get

[illegible]

[illegible] further repetitions we get an [illegible] set of [illegible] elements [illegible] of K

[illegible] and

[illegible] of degree [illegible]

[illegible]

then [illegible]

[illegible]

[illegible] was supposed to be an arbitrary [illegible]

Theorem [illegible] field containing the quotientfield Q of K [z]

$$[illegible] \qquad [5.2]$$

[illegible] element which [illegible]

As [illegible] that there is [illegible] or ϕ [illegible]

So we can apply [illegible] elements

[illegible]

of K in the [illegible] are the polynomials in z [See (1, 1), (1, 2)]

If we [illegible] are uniquely [illegible] of the elements [illegible] in z of degree [illegible] of every element [illegible]

Theorem [illegible] in one and only one manner by [illegible] polynomial in z, whose degree is > 0, for $i > 1$

The degree of a polynomial [illegible] function [illegible] from that [illegible] we get therefore the

[illegible] ments of Q and every element of Q is represented by a finite [illegible]

4

We will now use the representation of real numbers by continued fractions in order to approximate the elements [illegible] (5, 2).

[5/8] As (1, 15) holds,

$$x_1 - \frac{P_k}{Q_k} = \frac{(-1)^{k-1}}{x_2 \ \ldots} \ \frac{[illegible]}{Q_k}$$

we will prove that the right side of this equation is an element whose degree decreases indefinitely as k increases.

The elements x_i have been defined by x_i [illegible], a common factor being arbitrary (see (1, 12)) hence $x_2 \ldots x_{k+2} = x_2 \ldots$ [illegible]; therefore

$$\text{degree}\left(x_1 - \frac{P_k}{Q_k}\right) = -\text{degree}\ Q_k - d_2 - \ldots - [illegible] \tag{5, 5}$$

$d_i > 0$ for $k > 1$ and $x_i = a_i + 1/x_{i+1}$ (see 1, [illegible]). As the degree of a sum of summands of different degrees is equal to the highest of the degrees of the summands degree $(x_i) =$ degree $(a_i) = d_i$ for $i > 1$

$$Q_1 = 1,\ Q_2 = a_2,\ \ldots,\ Q_i = a_i\, Q_{i-1} + Q_{i-2}.$$

Hence degree $(Q_i) > 1$ for $i > 1$, hence the degrees increase with the index and therefore

$$\text{degree}\ Q_k = \text{degree}\ (a_k Q_{k-1}) = [illegible] + \text{degree}\ Q_{k-1} = \sum_{i=2}^{k} d_i \tag{5, 6}$$

Hence from (5, 5) and (5, 6) it follows that

$$\text{degree}\left(x_1 - \frac{P_k}{Q_k}\right) = -2 \sum_{i=2}^{k} d_i - d_{k+1} =$$

$$\text{degree}\left(\frac{-1}{Q_k Q_{k+1}}\right) \leq 1 - 2k. \tag{5, 7}$$

The efficiency of the approximation of an irrational number x by a continued fraction becomes clear by the theorem that if $\frac{a}{b}$ approximates x better than $\frac{P_k}{Q_k}$, then $b > Q_k$ holds. The corresponding theorem holds in the case we consider here.

Theorem. If $f(x)$ and $g(x)$ are polynomials of $K[x]$

$$\frac{f(x)}{g(x)} - \frac{P_n(x)}{Q_n(x)} \neq 0 \quad \text{and}$$

$$d' = \text{degree}\left(\alpha - \frac{f(x)}{g(x)}\right) < d = \text{degree}\left(\alpha - \frac{P_n}{Q_n}\right), \text{ then}$$

$$\text{degree } g(x) > \text{degree } Q_n \text{ holds} \tag{5. 8}$$

Proof. $\frac{f}{g} - \frac{P_n}{Q_n} = \left(\alpha - \frac{P_n}{Q_n}\right) - \left(\alpha - \frac{f}{g}\right)$. The two summands on the right side have different degrees, hence [illegible] degree [illegible] holds. As $(fQ_n - gP_n)$ is a polynomial, $0 \leq$ degree $(fQ_n - gP_n)$ = degree $g(x)$ + degree $Q_n + d$

degree $g(x) \geq -$ degree $Q_n - d > -$ degree $Q_n - d$ = degree Q_{n+1}

$>$ degree Q_n [see (5. 7), (5. 8)].

This theorem enables us to approximate functions given by a power series of $\frac{1}{x}$ by rational function in the neighbourhood of $x = 1/0$.

Exercise. $\frac{1}{2} \log \frac{x+1}{x-1} = x^{-1} + \frac{1}{3} x^{-3} + \frac{1}{5} x^{-5} + \ldots$

Represent this function by a continued fraction and approximate it by rational functions.

Lemma. Let $\alpha = (a_1, \ldots)$, $\alpha' = (a'_1, \ldots)$, let m be the lowest index [5/4]
for which $a_m \neq a'_m$ holds and $(a_1, \ldots, a_m) = A$, $(a'_1, \ldots, a'_m) = A'$ then

$$\text{degree } (\alpha - \alpha') = \text{degree } (A - A') \text{ holds} \tag{5. 9}$$

Proof. Without loss of generality we suppose that degree $a_m = r \geq$ degree $a'_m = r'$. We shall use the ordinary notations for the convergents of α and for those of α' we shall use them with a dash.

$P_i = P'_i$

$Q_i = Q'_i \quad$ for $i < m$

$Q_m = a_m Q_{m-1} + Q_{m-2} \qquad$ degree $Q_m = r + q$

$Q'_m = a'_m Q_{m-1} + Q_{m-2} \qquad$ degree $Q'_m = r' + q$

$$A - A' = \left(\frac{P'}{Q} [illegible]\right) - \left(\frac{P}{[illegible]} + \frac{(-1)^{n}}{Q_{n-1}\, Q}\right)$$

$$= (-1)^{n} \frac{x_n - x'_n}{Q_n\, Q'_n}.$$

1. If $r = r'$, degree $(x_n - x'_n) \geqq 0$

degree $(A - A')$ [illegible]

2. If $r > r'$, degree $(x_n - x'_n) = r$

[illegible] degree $\frac{1}{Q'_n{}^2}$. Hence

$$\text{degree } (A - A') \geqq \text{degree } \frac{1}{Q'_n{}^2} \qquad (5,\ 10)$$

holds in every case.

From (5, 10) it follows that

degree $(A$ [illegible] degree $\frac{1}{Q_{n}\,Q}$ [illegible] degree $\frac{1}{Q_n{}^2}$ [illegible] degree $\frac{1}{Q}$ [illegible]

and that degree $(A$ [illegible] $\frac{1}{Q\;\; Q_{n+1}}$ [illegible] degree $\frac{1}{Q}$ [illegible]. Hence

degree [illegible] degree $[A - A$ [illegible] A [illegible] A [illegible] degree $(A - A')$

[illegible] the degree of the first [illegible] of the [illegible] summands is greater [illegible] the degree [illegible] the two other summands.

Theorem [illegible] Let [illegible] x_2 [illegible] continued [illegible] polynomials of $K(x)$ and let [illegible] > 1 [illegible] degree x [illegible] $>$ [illegible] there exists a continued fraction $(x_1$ [illegible] x_2 [illegible]

For [illegible] the [illegible] x_1, x_2 [illegible] uniquely the values

P_1 [illegible] Q_1 [illegible] and P_N [illegible] $Q_N = (x_1$ [illegible] $x_N)$

Let $1 < $ [illegible] $< N$

From the preceding lemma [illegible] follows that [illegible] P_N [illegible] $Q_N - P_n\, Q_{n}$ [illegible]

$$= \text{degree } (P_{n+1}\;\; Q_{n+1} - P_n\;\; Q_n) = \text{degree} \left(\frac{1}{Q_n\, Q_{n+1}}\right) = - h_n.$$

where k_n increases to infinity with n

$$\frac{P_n}{Q_n} = \sum_{\nu=-m}^{\infty} a_\nu x^\nu = \sum_{\nu=-m}^{k_n} b_\nu x^\nu + \sum_{\nu=k_n+1}^{\infty} c_\nu x^\nu \qquad (5, 11)$$

The coefficients b_ν are independent of N. As k_n increases with the index n we get an infinite set $b_{-m}, \ldots, b_{k_n}, \ldots$ defining

$$\phi(x) = \sum_{\nu=-m}^{\infty} b_\nu x^\nu.$$

Finally we have to prove that $\phi(x) = (a_1, a_2, \ldots)$. Let $\phi(x) = (a'_1, a'_2, \ldots)$, and m be the smallest index for which $a_m \neq a'_m$; then it follows from the lemma that for every $n > m$,

degree $(\phi(x) - (a_1, \ldots, a_n))$ = degree $(\phi(x) - (a_1, \ldots, a_m))$ holds.

But $\phi(x) - (a_1, \ldots, a_n) = b_{k_n+1} x^{-k_n-1} + \ldots$ is of degree $-k_n - 1$ and decreases infinitely with n.

6. Continued Fractions with Rational Elements

Let R be the field of the rational numbers; then every finite continued fraction

$$(a_1, \ldots, a_n) = \frac{P_n}{Q_n} = \frac{P_1}{Q_1} + \left(\frac{P_2}{Q_2} - \frac{P_1}{Q_1}\right) + \cdots + \left(\frac{P_n}{Q_n} - \frac{P_{n-1}}{Q_{n-1}}\right) \qquad (6, 1)$$

$$= a_1 + \frac{1}{Q_1 Q_2} + \cdots + \frac{(-1)^n}{Q_{n-1} Q_n}$$

represents a rational number, but an infinite continued fraction defines a number if and only if

$$\sum \frac{(-1)^n}{Q_{n-1} Q_n} \qquad (6, 2)$$

converges. If the sum (6, 2) is convergent

$$(a_1, a_2, \ldots) \qquad (6, 3)$$

defines a real number equal to (6, 3). A necessary condition for the convergence of (6, 3) is therefore

$$|Q_{n-1}Q_n| \longrightarrow \infty \tag{6, 4}$$

If the numbers Q_n are either > 0 each, or < 0 each, or of alternating sign, the sum (6, 2) becomes an alternating sum. Hence the continued fraction converges if $|Q_n Q_{n-1}|$ increases steadily and (6, 4) is satisfied.

[6/2] *Theorem.* If $\sum |a_n|$ is convergent, (6, 3) is divergent.*

Proof. We prove by mathematical induction that

$$Q_n < \prod_{\nu=1}^{n} (1 + |a_\nu|) \tag{6, 5}$$

As $Q_1 = 1$, $Q_2 = a_2$, the formula holds for $n \le 2$. If (6, 5) is true for $n < m$,

$$Q_m = a_m Q_{m-1} + Q_{m-2}$$

$$|Q_m| \le \prod_{\nu=1}^{m-2} (1 + |a_\nu|)\,(|a_m|(1 + |a_{m-1}|) + 1)$$

$$< \prod_{\nu=1}^{m} (1 + |a_\nu|).$$

If $\sum a_n$ converges, the infinite product $\prod (1 + |a_\nu|)$ converges to a positive number Q and $|Q_n| < Q$ holds for every index n. Hence (6, 4) does not hold and the continued fraction is divergent.

[6/3] Let $a_n > 0$ for $n > 1$. Then as $Q_1 = 1$, $Q_2 = a_2 > 0$, $Q_n = a_n Q_{n-1} + Q_{n-2}$, it follows by mathematical induction that each number $Q_n > 0$.

$$Q_n Q_{n-1} = a_n Q_{n-1}^2 + Q_{n-1} Q_{n-2} > Q_{n-1} Q_{n-2} > 0$$

(6, 2) is therefore an alternating series whose elements have steadily decreasing absolute values. This series converges therefore if and only if (6, 4) is satisfied. These considerations lead to the following

Theorem. Let $a_n > 0$ for $n > 1$. Then the continued fraction (6, 3) is convergent if and only if $\sum a_n$ is divergent.

Proof. If $\sum a_n = \sum |a_n|$ is convergent, the continued fraction is divergent, as it has been proved by the preceding theorem.

* We use the term "divergent" for every non-convergent series.

Let $\sum a_i$ be divergent, then $a_i \to \infty$. As $Q_0 > 0$, $Q_1 = 1$, and $Q_{2n+1} = a_{2n+1} Q_{2n} + Q_{2n-1}$ we get by mathematical induction that $Q_{2n+1} \geq 1$. Hence $Q_{2n} = a_{2n} Q_{2n-1} + Q_{2n-2} \geq a_{2n} + Q_{2n-2}$ and as $Q_2 = a_2$, it follows by mathematical induction that

$$Q_{2n} \geq \sum_{r=1}^{n} a_{2r}$$

Hence

$$Q_{2n+1} Q_{2n} \to \infty, \quad \text{and} \quad Q_{2n} Q_{2n-1} \to \infty$$

therefore (6.4) is satisfied and as we stated above the condition is sufficient for the convergence of (6.3) in the case considered here.

Let $a_1 = 0$, $a_r \geq 1$ for $r > 1$, then $\sum a_r \to \infty$, and it follows from the (6.4)
preceding theorem that the continued fraction (6.5) converges to a value in the interval

$$\left(\frac{P_1}{Q_1}, \frac{P_2}{Q_2} \right) = (0, 1)$$

We will show that this value is *irrational*.

Let $x_1 = (a_1, a_2, \ldots)$ be rational, say $x_1 = \frac{d_2}{d_1}$ where d_1, d_2 are integral;

then $d_1 > d_2$ and $\frac{d_2}{d_1} = \frac{1}{a_2 + x_2}$, hence $x_2 = \frac{d_1}{d_2} - a_2 = \frac{d_3}{d_2}$, where $d_3 = d_1 - a_2 d_2$ is

integral as $x_2 = (0, a_3, \ldots)$, therefore $0 < x_2 < 1$, hence $d_2 > d_3 > 0$. In the same manner we get

$$x_2 = \frac{1}{a_3 + x_3}, \quad x_3 = [illegible] = \frac{d_4}{d_3} \quad \text{and } d_3 > d_4 > 0$$

By repetition of this procedure we get an infinite set of decreasing integral positive numbers

$$d_1 > d_2 > d_3 > d_4 > \ldots$$

and that is impossible.

Hence x_1 is irrational.

Example. $s_1 = 0$, $s_2 = 2$, $s_3 = 1$, and for $m > 1$

$$s_{2m} = \frac{2 \cdot 4 \cdots 2m-2}{1 \cdots 2m-1} > 1, \quad s_{2m+1} = \frac{3 \cdot 5 \cdots 2m-1}{2 \cdot 4 \cdots 2m-2} > 1;$$

then it is

$$Q_1 = 1, \; Q_2 = 2, \; Q_3 = 3$$

and from the identities

$$2 \cdot 4 \cdots 2m = 1 \cdot 3 \cdots (2m-1) \, s_{2m} + 2 \cdot 4 \cdots 2m-2,$$

$$1 \cdot 3 \cdots 2m+1 = 2 \cdot 4 \cdots 2m \, . \, s_{2m+1} + 1 \cdot 3 \cdots 2m-1,$$

it follows by mathematical induction that

$$Q_{2m} = 2 \cdot 4 \cdots 2m, \quad Q_{2m+1} = 1 \cdot 3 \cdots 2m+1,$$

hence

$$Q_n \, . \, Q_{n-1} = n!;$$

the continued fraction is irrational. Its value is

$$\frac{1}{Q_1 Q_2} - \frac{1}{Q_2 Q_3} + \cdots + \frac{(-1)^n}{Q_{n-1} Q_n} + \cdots = \sum (-1)^n / n! = e^{-1}.$$

Hence e is irrational.

PART IV.

APPROXIMATE SOLUTION

§ 1. Horner's scheme

Let K be an arbitrary field, e.g., the field of the real numbers or (1/1) of the complex numbers, q be an element of K and $f(x)$ a polynomial of $\mathrm{K}[x]$.

$$f(x) = \sum_{i=0}^{n} a_i x^i = (x-q)\sum_{i=1}^{n} a'_i x^{i-1} + a'_0 = (x-q) f_1(x) + a'_0$$

then $a_i = a'_i - q a'_{i+1}$. Hence

$$a'_n = a_n$$

$$a'_{k} = a_{k} + q a'_{k+1} \qquad (1.1)$$

[illegible]

$$a'_0 = f(q)$$

holds. We can arrange the calculation of the coefficients a'_i as follows:

$$\begin{array}{cccccc} a_n & a_{n-1} & a_{n-2} & \dots & a_1 & a_0 \\ & q a'_n & q a'_{n-1} & \dots & q a'_2 & q a'_1 \\ a'_n & a'_n & a'_{n-2} & \dots & a'_1 & a'_0 \end{array} \qquad (1.1')$$

$$f(q) = a'_0 \qquad (1.2)$$

We can find now by the same method $f_1(x)$ satisfying

$$f_1(x) = (x-q) f_{11}(x) + a''_1, \quad f_{11}(x) = \sum_{i=2}^{n} a''_i x^{i-2}$$

After $n-1$ steps we get $f(x)$ represented as a polynomial in $x-q$:

$$f(x) = a'_0 + a''_1(x-q) + \dots + a^{(n)}_{n-1}(x-q)^{n-1} + a_n(x-q)^n$$

This representation is known in Analysis as the *Taylor series* of $f(x)$ at the point $x = q$. The successive calculation of the coefficients can easily be done on using the last line of (1.1') up to a'_1. The complete scheme for this calculation is called *Horner's scheme*. It is the most convenient method

for calculating $f(q)$ if q and the coefficients of $f(x)$ are given, and furnishes the representation of $f(x)$ by $x-q$. The calculation will be explained by the following

Example. $f(x)=x^4-15x^3+68x^2-119x+67$

$q=1.$	1	-15	68	-119	67
		1	-14	54	-65
	1	-14	54	-65	2
		1	-13	41	
	1	-13	41	-24	
		1	-12		
	1	-12	29		
		1			
	1	-11			

$$f(x)=(x^3-14x^2+54x-65)(x-1)+2$$
$$=(x^2-13x+41)(x-1)^2-24(x-1)+2$$
$$=(x-12)(x-1)^3+29(x-1)^2-24(x-1)+2$$
$$=(x-1)^4-11(x-1)^3+29(x-1)^2-24(x-1)+2.$$

[1,2] Horner's scheme is very useful for calculating the roots. The method will be explained by the help of the above example.

$x=y+1$, $f(x)=g(y)=y^4-11y^3+29y^2-24y+2$

$$f(1)=g(0)=2$$
$$f'(1)=g'(0)=-24$$

We therefore suspect that there is a root of $f(x)$ near $x=1$. In the neighbourhood* of $x=1$, $g(y)\backsim-24y+2$, and $-24y+2=0$ for $y\backsim0.1$. Therefore we represent $g(y)$ by a polynomial in $y-0.1$.

$y=0.1$	1	-11	29	-24	2
		$+0.1$	-1.09	$+2.791$	-2.1209
	1	-10.9	27.91	-21.209	-0.1209
		$+0.1$	-1.08	$+2.683$	
	1	-10.8	26.83	-18.526	
		$+0.1$	-1.07		
	1	-10.7	25.76		
		$+0.1$			
	1	-10.6			

* The sign $\backsim$ means "approximately equal to". We apply here an elementary theorem on continuous functions which we proved in [illegible] 2.2.

As $f(1.1) =$ [illegible] there is a root between 1.1 and 1.2. We approximate therefore $f(x)$ by

$18.326\,(x-1.1)-0.1209$, hence $x-1.1 \approx 0.007$

$q =$ 0.007	1	[illegible]	[illegible]	[illegible]	[illegible]
		0.007	[illegible]	[illegible]	[illegible]
	1	10.007	25.[illegible]	−18.70[illegible]	0.00104787894
		− 0.007	+ 0.07[illegible]	− 0.18[illegible]	
	1	10.614	[illegible]	18.88[illegible]	
		− 0.007	+ 0.074347		
	1	− 10.621	25.092894		
		0.007			
	1	− 10.628			

Hence the root is approximately equal to [illegible]. The next approximation is q [illegible]. If we continue in exactly the same manner the calculation would become very laborious. We therefore neglect the terms q^3 and q^4 and we get a good approximation by taking terms up to q^2. For this approximation we are led to the following equation [illegible] q^2 [illegible] q^4 [illegible]. As $q \sim$ [illegible] 10^{-4} the two last terms will influence only the [illegible] and the following decimals of the equation; for [illegible] $10^{-4} < q <$ [illegible] 10^{-4} the quadratic term becomes 0.000007...

Hence $q =$ [illegible] $x =$ [illegible]

On using this approximation we can easily get some more figures of the decimal development. The most difficult task is sometimes to get a first approximation of the roots. For this purpose it is often helpful to know the values of $f(x)$ for a suitable set of values x. We may get these values by Horner's scheme, but it is often useful to abbreviate this scheme in the following manner.

Given $q_1, q_2, \ldots q_n$ [illegible]. We calculate by Horner's scheme [illegible]

$$f(x)=b_1+(x-q_1)\,f_1(x) \qquad f(q_1)=b_1$$

$$f_1(x)=b_2+(x-q_2)\,f_2(x) \qquad f(q_2)=b_1+(q_2-q_1)\,b_2$$

$$\ldots \qquad \ldots$$

$$f_{n-1}(x)=b_n+(x-q_n)\,f_n(x)$$

[illegible]

We will [illegible] the [illegible] used [illegible] in the previous example in this manner.

$q_1 = 1$	1	-15	68	-119	67
		1	-14	54	-65
$q_2 = 2$					
	1	-14	[illegible]	[illegible]	2
		2	[illegible]	[illegible]	
$q_3 = [illegible]$					
	1	-12	[illegible]		
		[illegible]	[illegible]		
	1	[illegible]	[illegible]		

$f(x) = [illegible] (x-2)(x-3)(x-4)$,

$f(0) = 67$, $f(1) = 2$, [illegible] $+17$, [illegible]

This representation shows that for $x < 1$ [illegible] each of the terms becomes > 0 and that for $x > [illegible]$ $x > 0$, [illegible] the 1st, the [illegible] and the sum of the two other terms become > 0. [illegible] roots are situated in the interval [illegible]. We can [illegible] one root in the interval [illegible] 2 [illegible] at least one root in the interval [illegible] $7 + 7 [illegible] - [illegible] - 117$. Hence there is a root in the interval [illegible]. The reader may calculate it by Horner's scheme as an *exercise*.

[1.4] We will use here a different method of calculation. If $f(x) = \sum_{0}^{n} a_\nu x^\nu = [illegible]$, $\frac{1}{x}$ satisfies the condition $\sum_{0}^{n} a_{n-\nu} [illegible] \left(\frac{1}{x}\right) [illegible] = 0$, and to every root of x in the interval [illegible] there corresponds a value of $\frac{1}{x} >$ [illegible]. Thus we can [illegible] and to the following method of approximation due to *Lagrange*.

If ξ is a root of $g(x)$, $a < \xi < a + 1$ [illegible] $\frac{1}{[illegible]}$ [illegible] $= [illegible]$ $g_1\left(\frac{1}{x}\right)$, $q_1 > 1$ is a root of [illegible] and $b \leq q_1 < b + 1$ [illegible] $q_1 = b + \frac{1}{q_2}$. By repetition of the procedure we [illegible] get a representation of ξ as a continued fraction. If we stop the calculation after n steps we get the approximation $\frac{P_n}{Q_n}$ and the error becomes

$$\left|\xi - \frac{P_n}{Q_n}\right| < \frac{1}{[illegible] Q_{n-1}} < \frac{1}{Q_n^2}$$

This method will be illustrated by the example previously used. We know that $x^4 - 15x^3 + 68x^2 - 119x + 67$ has a root in the interval (8, 9). Therefore we represent this polynomial by Horner's method as a polynomial in $x - 8$.

$q=8$	1	-15	68	-119	67
		8	-56	96	-184
	1	-7	12	-23	-117
		8	8	160	
	1	1	20	137	
		8	72		
	1	9	92		
		8			
	1	17			

Hence $117\, q_1^4 - 137\, q_1^3 - 92\, q_1^2 - 17\, q_1 - 1 = 0$.

By means of an analogous scheme we see that for $q = 2$ the last coefficient becomes positive, but for $q = 1$ it becomes negative, so q_1 is in the interval (1, 2). Thus we have to make the Horner development for $q = 1$.

$q=1$	117	-137	-92	-17	-1
		117	-20	-112	-129
	117	-20	-112	-129	-130
		117	97	-15	
	117	97	-15	-144	
		117	214		
	117	214	199		
		117			
	117	331			

In the same manner as it has been done for q_1 we see that q_2 is situated in the interval (1, 2).

$q=1$	130	144	-199	-331	-117
		130	274	75	-256
	130	274	75	-256	-373
		130	404	479	
	130	404	479	223	
		130	534		
	130	534	1013		
		130			
	130	664			

η_3 is in the interval (2, 3)

$q = 2$	378	-223	-1018	$+604$	-130
		[illegible]	1046	66	-1160
		[illegible]	[illegible]	[illegible]	[illegible]
		[illegible]	2548	5142	
	773	1299	2571	4744	
		[illegible]	[illegible]		
	174	2013	[illegible]		
		[illegible]			
	173	[illegible]			

[illegible] the reader has [illegible] experience that q has to be chosen so that the sign of the last coefficient does not change, but that the procedure [illegible] for $q = 1$ [illegible] the sign [illegible] $q = 1$ will alter the sign of the [illegible] in the [illegible] row [illegible] scheme that the third coefficient will not change its sign, [illegible] being too big. Hence the [illegible] coefficients [illegible] and therefore [illegible] become negative. However [illegible] counterbalanced by more than [illegible] and therefore [illegible] by more than [illegible] and [illegible] of the last coefficient would be altered. Hence $q = 1$

$q = 1$	1326	-4311	[illegible]	[illegible]	[illegible]
		[illegible]	[illegible]	-11244	[illegible]
	1326	700	[illegible]	[illegible]	[illegible]
		[illegible]	[illegible]	[illegible]	
	1326	[illegible]	[illegible]	[illegible]	
		[illegible]	[illegible]		
	1326	[illegible]	[illegible]		
		[illegible]			
	1326	[illegible]			

The next q will become $= 1$. Hence $\xi = (8, 1, 1, 2, 4, 1, \ldots)$

$P_1 = 8$ $Q_1 = 1$

$P_2 = 9$ $Q_2 = 1$

$P_3 = 17$ $Q_3 = 2$

$P_4 = 43$ $Q_4 = 5$

$P_5 = 189$ $Q_5 = 22$

$P_6 = 232$ $Q_6 = 27$

$Q_7 \geq 49$

[illegible] the error [illegible] is [illegible] and [illegible]

As [illegible] second decimal only the third decimal may [illegible] the reader [illegible] see that Lagrange's method is [illegible] for practical calculation.

In the last subsections Horner's scheme [illegible] for the solution of equations with real coefficients by real roots. [illegible] can be applied [illegible] has been stated in the beginning [illegible] arbitrary fields. We will use [illegible] to find out [illegible] complex numbers. Let [illegible] be [illegible] polynomial [1/5]

$$b_0 + b_1 + b_2 \quad [illegible]$$

$$a b_1 + (a + a^2) b_2 \quad [illegible]$$

[illegible]

Hence if a is a root, $\sum b_i = \sum b_i \, a$ [illegible]

Let b_i be positive numbers and [illegible].

1) If $|a| < 1$ the equation [illegible]

2) If $|a| = 1$, $a = \cos\theta + i\sin\theta$,

Σb_i [illegible] Σb_i [illegible] but in this case the last coefficient [illegible] Σ [illegible] $>$ [illegible]
If therefore in $a_0 x^n + a_1 x^{n-1} + \dots$

$$a_0 < a_1 < \dots < 1, \qquad [illegible]$$

then for every root [illegible] of the polynomial [illegible] Hence [illegible] the roots [illegible] of $\Sigma a_i x^i$ satisfy [illegible] < 1. This theorem is known as

[illegible] theorem: The complex roots of $\Sigma a_i x^i$ have absolute values < 1 if the coefficients satisfy [illegible]

6

§ 2. THE ROOTS OF REAL POLYNOMIALS

[2/1] In this section

$$a, \quad b, \quad c, \quad d, \quad e \qquad (2,1)$$

[illegible] real numbers [illegible] the same manner

$$\alpha, \quad \beta, \quad \gamma, \quad \delta \qquad (2,2)$$

[illegible] numbers [illegible]

Hence $a+\bar{a}$ is real; $a\bar{a}$ is positive, [illegible] and if [illegible] $= 0$, a is real.

$$f(x) = a_0 + a_1 x + \dots + a_{n-1} x^{n-1} + x^n \qquad (2,3)$$

can be represented by $f(x) = \prod_{i=1}^{n} (x - \alpha_i)$ (2,4)

(see Part II [18 2]).

Theorem. If α is a root of (2,4), $\bar{\alpha}$ is also a root of (2,4).

1st Proof. Let K be the field of real numbers, i and $-i$ be the roots of x^2+1, then $K(i) = K(-i)$ is the field of the complex numbers and there is an automorphism J of this field interchanging i with $-i$ and leaving the real numbers unaltered. $f(x)$ will not be altered by J, hence α will be transformed into a root of $f(x)$, but as α will be transformed into $\bar{\alpha}$, the theorem is true.

2nd Proof. If α is real, $\bar{\alpha} = \alpha$. If α is not real $(x-\bar{\alpha})(x-\alpha) = g(x)$ is a real polynomial and irreducible in the field of the real numbers. As $f(x)$ and $g(x)$ have a common root these polynomials have a common factor of positive degree. Hence $f(x)$ is divisible by $g(x)$ and $\bar{\alpha}$ is therefore a root of $f(x)$.

Corollary 1.

$$f(x) = (x - c_1) \dots (x - c_r)(x - \beta_1)(x - \bar{\beta}_1) \dots (x - \beta_k)(x - \bar{\beta}_k) \qquad (2,5)$$

where $n = r + 2k$.

Corollary [illegible] If every real root [illegible] is counted as many times as its order of multiplicity [illegible] the number of the real roots [illegible] are both even or both odd.

Corollary 3. If n is odd, there exists at least one real root.

$$\Delta = \prod_{i<j} (\alpha_i - \alpha_j)^2 \tag{2.6}$$

is called the discriminant of $f(x)$. As Δ is a symmetric polynomial in $\alpha_1, \ldots, \alpha_n$ with integral coefficients it follows (see Part II [illegible]) that

$$\Delta = g(a_1, \ldots, a_n) \tag{2.7}$$

where g is a polynomial with integral coefficients. From (2.6) it follows (see Part II, [10/8]) that

$$\Delta = 0 \quad \text{if and only if} \quad \alpha_i = \alpha_j \text{ for } i \neq j. \tag{2.8}$$

Let the n roots α_i be all different. As it has been proved in Part II [illegible],

$$\Delta = \delta^2, \quad \delta = \begin{vmatrix} \alpha_1^{n-1} & \cdots & \alpha_1 & 1 \\ \vdots & & \vdots & \vdots \\ \alpha_n^{n-1} & \cdots & \alpha_n & 1 \end{vmatrix} \tag{2.9}$$

To get $\bar{\delta}$ we have to interchange every root with its conjugate. From [illegible] it follows that this permutation [illegible] interchanges of rows in the determinant (2.9).

Hence $\bar{\delta} = (-1)^k \delta$, and therefore

$$(-1)^k \Delta = (-1)^k \delta^2 = \delta\bar{\delta} > 0$$

Hence the following theorem holds.

Theorem. Let $f(x)$ of degree n have n different roots. Then the discriminant of $f(x)$ is positive (negative) when the number of pairs of conjugate non-real roots is even (odd).

Corollary. A real polynomial of degree 3 has three real roots if and only if the discriminant is positive.

A real polynomial of degree 4 with positive discriminant has either four different real roots or two pairs of conjugate complex roots.

Exercise. Prove the preceding theorem without the help of (2.8).

[2,2] If we [illegible] continuous and differentiable function. Hence

1. If $a < b$ and the [illegible] different, then there is a root of $f(x)$ in the interval (a, b).

2. If [illegible] of [illegible] in the interval (a, b).

3. If $f(x) = (x-a)^k g(x)$, $g(a) \neq 0$, $k > 0$, then

$$f'(x) = (x-a)^{k-1} g_1(x), \quad g_1(a) \neq 0$$

For, $g_1 [illegible] k\,g(x) + (x-a)\,g'(x)$

[illegible]

[illegible] with a fin[illegible] number [illegible] If [illegible] coefficients of [illegible] positive [illegible] Hence $f(x)$ is not zero for any positive value of x. Thus

[illegible] in the sequence of [illegible] So we [illegible] the connection between the [illegible] of [illegible] coefficients. The experience we [illegible] very helpful to us.

By developing [illegible] polynomial in $x-b$, we get

$$f(x) = f_b(x-b) = a_{b,0}(x-b)^n + \dots + a_{b,n} \tag{2, 10}$$

So [illegible] every real number b and th[illegible] there belongs a set of $n+1$ real numbers [illegible]

such that [illegible] independent of b (2, 11)

[illegible]

Hence [illegible] and for [illegible]

$$c_{k,n} = \frac{1}{k!} f^k(b). \tag{2, 12}$$

The sequence

$$c_{0,n},\ c_{1,n},\ \dots,\ c_{n,n} \tag{2, 13}$$

may be reduced by [illegible] every element [illegible] 0. If in the reduced sequence [illegible] has a sign different from the sign of the preceding element, it [illegible]. We [illegible] the number [illegible] of these changes

$$0 \leq C(b) \leq n \tag{2, 14}$$

If the first and the last element of the [illegible] have the same sign, [illegible] if they have different signs [illegible] (2, 11) it follows therefore that [illegible] where [illegible] and it is odd where $a_n f(b) < 0$.

If we strike out [illegible] element which [illegible] different from the first, the number of changes [illegible], if we strike out an arbitrary element, the number of changes [illegible] the [illegible] We will compare the changes in the rows

$$\begin{matrix} [illegible] \\ [illegible] \end{matrix}$$

where [illegible]

Let [illegible] > 0. If [illegible] change in the 2^d row [illegible] 0 and it is of the same sign as [illegible]. Hence the 2^d row has the same number of changes as the subsequence formed by those elements [illegible] of the first row which have the same sign as the corresponding elements of the second row. The number of changes of the 2^d row is therefore less or equal to the number of the changes in the first. The same holds for every pair of subsequent rows in the following system which we get by Horner's scheme

$$\begin{matrix}
a_n & a_{n-1} & \dots & a_2 & a_1 & a_0 \\
a_n & a'_{n-1} & \dots & a'_2 & a'_1 & [illegible] \\
a_n & a''_{n-1} & \dots & a''_2 & [illegible] & \\
[illegible] & [illegible] & \dots & [illegible] & [illegible] & [illegible] \\
 & & & & & \\
[illegible] & [illegible] & \dots & [illegible] & [illegible] & [illegible]
\end{matrix}$$

But the last row is identical with

$$[illegible]$$

Hence for $q > 0$ [illegible] $\leq$ [illegible] (2,15)

Let $c = b + q > b$

$f(x) = f_b(x - b) = f_c([x - b] - q)$

then it follows from (2,15) that the number of changes in f_c is not greater than the number of changes in f_b, i.e.

$$C(c) \leq C(b) \qquad (2,16)$$

As $f(x)$ and [illegible] have the same number of changes, we can suppose without any loss of generality that $a_0 > 0$. If a_0, [illegible], and q are positive $a_{q(n-1)} > a_0$ [illegible] and if q is great enough [illegible] become positive. In this manner we see that from a certain value of q the coefficients become all positive. Hence $C(q) = 0$ for $q > c$. Therefore

Theorem. C decreases steadily to 0.

Let $b < c$ and let these values be not separated by a root of $f(x)$, then a_n of f_b and a_n of f_c have the same sign. Hence $C(b) = C(c) + 2k$, where $k \geq 0$ is an integral number.

C is therefore a discontinuous decreasing function taking only integral values and the values at points which are not roots have even values. We have now to investigate the action at the roots of $f(x)$.

Let c be a root of multiplicity m, and $x - c = y$, then

$f(x) = [illegible]\, y^n + \dots + a_{n-m}\, y^m$, $a_{n-m} \neq 0$.

Let [illegible]

$$f(x) = f_e(x) = [illegible] + \dots + a_{n-m}\, x^m + \dots + a_n \qquad (2,17)$$

Let $a_{n-m}\, y^m = a_{n-m}(x + e)^m = g(x)$

If $e > 0$ there is no change in the coefficients of $g(x)$.

If $e < 0$, the coefficients have alternating signs and have therefore m changes. If e is small enough the last $m + 1$ coefficients of $f_e(x)$ have the same signs as $g(x)$. If therefore e increases from small negative to small positive values, the number of the changes in $a_{n-m}, \dots, a_n$ decreases by an

even [illegible] and the number of the changes [illegible] decrease by m.

Hence [illegible] $2 l + m$, where $l \geq 0$ is an integral number. By the help of these [illegible] we get the following theorem

Theorem [illegible] let r be the number of the roots of [illegible] in the interval [illegible] every [illegible] being counted with its own multiplicity, then

$$C(b) = C(c) + r + 2k, \qquad (2.16)$$

where $k \geq 0$ is an integral number.

Applying [illegible] that $C(c) = 0$, we get as corollary

Descartes' rule. The number of the positive roots of [illegible] being counted with its own multiplicity is equal to the number of the changes of signs of the coefficients of [illegible] to a number less than it by an even number.

If we consider that in (2.12) a_{n-s} and f^{s} [illegible] have the same sign, (2.16) can be expressed in the following manner

Budan-Fourier's theorem. [illegible] the number of the roots of [illegible] in the interval [illegible] every root being counted with its own multiplicity is equal to the difference of the changes of signs in the sets

$$f(b),\ f'(b), \ldots, f^{(n)}(b) \text{ and}$$

[illegible] or [illegible] even number [illegible]

If we set $x = \frac{[illegible] + b}{y + 1}$ and therefore $y = \frac{x - b}{[illegible] - x}$ then the positive roots of [illegible] correspond [illegible] to the roots of [illegible] in the interval [illegible] Therefore we are able to apply Descartes' rule to find out the number of the roots in this interval.

These formulae [illegible] give directly the exact number of the roots in an interval, but they are very useful for getting it even in more complicated cases.

We will go into further details of the example considered in [illegible]

$$f(x) = x^{4} - 15x^{3} + 68x^{2} - 119x + 67.$$

As we stated before, the [illegible] roots [illegible] positive and situated in the interval [illegible] We [illegible] get this result [illegible] considering the [illegible] of signs [illegible] has no change and [illegible] positive root [illegible] has no [illegible] root. [illegible] the [illegible] for this example we get—on considering the signs only—

$$C(0) = C(1) = 4$$
$$C(2) = 3$$
$$C(8) = 1$$

We know [illegible] roots, [illegible] the interval [illegible] another [illegible] the interval [illegible] each [illegible] there [illegible] change. We should find out [illegible] the [illegible] change [illegible] corresponds to [illegible] roots of [illegible] For this purpose we try to approximate these [illegible] roots by Horner's scheme and [illegible] very simple calculations

$$C(3) = 1 \qquad C(2\cdot 6) = 3$$
$$C(2\cdot 65) = 1 \qquad C(2\cdot 64) = 3$$

Hence the two roots [illegible] be situated in the interval $2\cdot 64 < x < 2\cdot 65$ but we [illegible] prove that [illegible] is negative [illegible] this interval. We stated previously that

[illegible]

[illegible] Hence for [illegible]

[illegible]

Hence [illegible] has [illegible] the two roots [illegible] The same result [illegible] obtained [illegible] calculating [illegible] and stating that it is negative.

[2/8] A method to [illegible] of the different roots [illegible] has been given by Sturm. We suppose that [illegible] and that therefore [illegible] has only single roots. [illegible] of [illegible] roots, [illegible] always [illegible] by the algorithms and if [illegible] of positive degree, [illegible] has to be replaced by [illegible] which has the same roots but only simple ones.

The method uses a [illegible] *Sturm* [illegible] of polynomials

$$f(x) = f_1(x), f_2(x), \ldots, f_m(x). \qquad (2, 10)$$

and the number $C(b)$ of the changes of sign in

$$f_1(b), \dots, f_m(b).$$

The chain should be made in such a manner that $C(x)$ is a steadily decreasing function changing its values only at the roots of $f(x)$ and having at these points a saltus of the value 1. Then

$$C(a) - C(b) \quad \text{for } a < b,\ f(a) \neq 0 \neq f(b) \tag{2.29}$$

becomes the number of the roots of $f(x)$ in the interval (a, b).

For this purpose we have to arrange the chain so that at each root of $f(x)$ one change will be lost and that in the roots of $f_\nu(x)$, $\nu > 1$, the number of the changes will not be altered. In order to get a loss of changes in the roots of $f(x)$, f_1 must take the sign of f_2 when x passes a root of $f(x)$ from the left to the right, i.e.

(1) $f_2(x)$ has the same sign as $f'(x)$.

In order to avoid an alteration of the number of changes at the roots of f_ν, $\nu > 1$,

(2) $f_m(x)$ should have constant sign, and

(3) for every root b of $f_{\nu+1}(x)$ we have $f_\nu(b) f_{\nu+2}(b) < 0$.

So $f_{\nu+1}(x)$ will have the sign of exactly one of its neighbours before and after passing a root. The outcome of these considerations is the following theorem.

Sturm's theorem. If the chain (2.28) satisfies the conditions 1–3, the number of the roots of $f(x)$ in any interval (a, b) is given by (2.29).

In order to get a chain of this kind we may use the algorithm to find the h. c. f.

$$f_2(x) = f'(x),\quad f_\nu(x) = q_\nu(x) f_{\nu+1}(x) - f_{\nu+2}(x),$$

where the degrees of the [illegible]

The first two conditions are obviously satisfied. As every common factor of $f_\nu(x)$ and $f_{\nu+1}(x)$ must be a common factor of $f_1(x) = f(x)$ and of $f_2(x) = f'(x)$, and as $f(x)$ and $f'(x)$ have no common factor, $f_\nu(x)$ and $f_{\nu+1}(x)$ have no common root. If therefore $f_{\nu+1}(b) = 0$, $f_\nu(b) = -f_{\nu+2}(b) \neq 0$. Hence

$$f_\nu(b) \cdot f_{\nu+2}(b) < 0.$$

By Sturm's method it is always possible to get the exact number of the different roots in any interval but the practical computation is sometimes very burdensome and it is often more convenient to use Budan Fourier's theorem in connection with special considerations as we did it in the previous example. The reader may investigate the same example with Sturm's method as an *exercise*.

Remark. Sturm's chain (2, 19) can always be replaced by $c_0 f_0, c_1 f_1, \ldots, c_n f_n$ where $c_0, c_1, \ldots, c_n$ are positive constants.

[illegible] Show that the [illegible] *Legendre's* polynomials

$$P_m(x) = \frac{1}{2^m m!} D^m[(x^2-1)^m]; \quad m = 0, 1, 2, \ldots \tag{2, 21}$$

D^m denotes the m-th derivative of the expression written in [ ] and D^0 is the function itself. If u and v are polynomials in x

$$D^m(uv) = \sum_{i=0}^{m} \binom{m}{i} D^{m-i}u \, D^i v \tag{2, 22}$$

$$D^m[(x^2-1)^m] = D^{m-1}[D(x^2-1)^m] = D^{m-1}[(x^2-1)^{m-1} 2mx],$$

whence from (2, 22) it follows for $m \geq 1$

$$D^m[(x^2-1)^m] = 2mxD^{m-1}[(x^2-1)^{m-1}] + 2m(m-1)D^{m-2}[(x^2-1)^{m-1}] \tag{2, 23}$$

On the other hand we get from (2, 22) for $m \geq 1$

$$2D^m[(x^2-1)^m] = 2D^m[(x^2-1)^{m-1}(x^2-1)] = 2\{(x^2-1)D^m(x^2-1)^{m-1}\}$$
$$+ 4mxD^{m-1}(x^2-1)^{m-1} + 2m(m-1)D^{m-2}(x^2-1)^{m-1} \tag{2, 24}$$

By subtraction of (2, 23) from (2, 24) and applying (2, 21) we get

$$mP_m(x) = (x^2-1)P'_{m-1}(x) + mxP_{m-1}(x) \tag{2, 25}$$

As $P_0(x) = 1$, $P_1(x) = x$, $P'_0(x) = 0$, (2, 25) holds also for $m = 1$ and therefore generally.

From

$$D^{m+1}[(x^2-1)^m] = D^m[2mx(x^2-1)^{m-1}]$$
$$= 2mxD^m[(x^2-1)^{m-1}] + 2m^2 D^{m-1}[(x^2-1)^{m-1}]$$

we get

$$P'_m(x) = xP'_{m-1}(x) + mP_{m-1}(x). \tag{2, 26}$$

From (2, 25) and (2, 26) we eliminate P'_{m-1} and get

$$(x^2-1)P'_m(x) = mxP_m(x) - mP_{m-1}(x). \tag{2, 27}$$

On replacing m by $m+1$ in (2, 25) we get

$$(m+1)[illegible]_{1}(x)=(x^2-1)P'_m(x)+(m+1)xP_m(x). \qquad (2,\ 25')$$

From (2, 27) and [illegible]) we eliminate $x\,P'_m(x)$ and we get

$$(m+1)P_{m+1}(x)=(2m+1)xP_m(x)-mP_{m-1}(x). \qquad (2,\ 28)$$

We consider the sequence

$$P_m(x),\ P_{m-1}(x),\ldots,P_1(x),\ P_0(x)=1 \qquad (2,\ 29)$$

in the interval

$$-1\leq x\leq +1$$

From [illegible] that [illegible] P_m [illegible] If P_m [illegible] and [illegible] have a common [illegible] real [illegible] of P_{m+1} [illegible] from (2, 28) and [illegible] of all subsequent [illegible] Hence for P_m [illegible] and [illegible] follows from (2, 28) that $P_{m+1}(x)P_{m-1}(x)<0$ at every root of $P_m(x)$.

[illegible] for every m.

From (2, 27) it follows that

$$P_m(1)=P_{m-1}(1)$$

and

$$[illegible]$$

As $P_0(x)=1$, it follows that

$$P_m(1)=1$$

and

$$P_m[illegible]$$

The [illegible] of changes of sign [illegible] is therefore [illegible] Hence [illegible] different roots in the interval [illegible] the roots of Legendre's [illegible] are all situated in the interval $-1, +1$ and are simple roots.

To find out systematically the real roots of a polynomial [illegible]

$$f(x)=a_0+a_1x+\ldots+a_nx^n$$

with real coefficients we have at first to find an interval containing all these roots.

Let

$$t = 1 + \left| \frac{a_{n-1}}{a_n} \right| + \cdots + \frac{a_1}{a_n} + \frac{a_0}{a_n}$$

As [illegible] $f(x)$ has the same sign as a_n [illegible] therefore [illegible] has the same [illegible]. Hence the roots of $f(x)$ are situated in the interval $(-t, +t)$.

When an interval containing all the real roots of the polynomial has been found [illegible] Sturm's method [illegible] how many roots are [illegible] intervals. The [illegible] intervals containing roots should [illegible] and the procedure must be repeated [illegible] an interval [illegible] found for every root [illegible] that $a < b <$ [illegible] and $b - a$ is [illegible] than the approximation required for the special problem. As the root [illegible] represented by a decimal fraction, the first subdivision [illegible] the second subdivision by [illegible] $m+2$ [illegible] the root will be determined up to the mth decimal.

[2, 6] This method to [illegible] roots [illegible] be used by a [illegible]. By the practice of [illegible] what [illegible] interval may [illegible] the polynomial only for the endpoints of [illegible] interval [illegible] containing only one simple root, the approximation [illegible] very quickly by Horner's scheme.

[illegible] If [illegible] have different signs, the graph of the function $f(x)$ may be replaced by the straight line connecting the points with abscissae a and b. This line intersects the x-axis in [illegible]. This value may be considered as a first approximation of the root. Let us consider the example of [illegible]

$$x = 1 + y, \quad f(x) = g(y) = y^4 - [illegible] y^3 - [illegible] y^2 - 21y + 2$$

$$g(0) = 2, \quad g(1) = -3.$$

Hence we get by the [illegible] an approximation y_1 [illegible] We will try to improve this approximation by

$$24 \ldots y_1^4 - [illegible] y_1^3 + 29 y_1^2 + 2 \backsim 0 \qquad (2.30)$$

Hence

$$y_2 \backsim 1.4.$$

This approximation is better, but it is not good. As we can calculate [illegible]

$$x - 1 = 0.0035324$$

Of course the graph of the polynomial is in that interval very different from a straight line. Now

$$g(0) = 2 \qquad g(1) = [illegible]$$

$$g'(0) = -24 \qquad g'(1) = [illegible]$$

The graph is therefore [illegible] in the interval and the root must obviously lie near the point $x - 1 = 0$.

Newton's method. The graph is replaced approximately by the tangent, i.e. the terms of higher degree in Horner's scheme have to be omitted. This method gives good results when the distance from the root is small, the term of first degree has a coefficient of high absolute value and the absolute values of the coefficients of the higher terms are not too big. This method has been used in the above example, and gives the approximation $2 - 24y_1$ [illegible]. We can improve the solution by using the equation (3), with $y_1 = 0.1$ then y_2 becomes < 0.065.

As for $\qquad y_1 = 0, \qquad y_2 > 0.082$

and for $\qquad y_1 = 0.1, \qquad y_2 < 0.065$ hold,

and as y_2 is a continuous function of y_1 there must be in the interval (0.082, 0.065) a value for which $y_1 = y_2$, and that is a root. So Newton's method is very useful in certain cases, but if the interval is big or the tangent makes only a small angle with the axis the method cannot be used.

The suitable choice of the methods should be learned by practice. In this section sometimes reference has been made to the graph of a polynomial. So the reader may ask if *graphical methods* may not be helpful to get the roots of a polynomial. Of course methods of this kind exist and are very helpful to get a convenient first approximation for the roots, but in applying these methods only those readers may succeed who are familiar with the theory and practice of mathematical drawing.*

Let c be a fixed real number $\neq 0$ and let b and [illegible] be two different and consecutive roots of $f(x)$, then $f'(x)$ has a constant sign in the interval [2.7]

* A very useful graphical method is e.g. the rectangular method. For reference see Bieberbach-Bauer, Vorlesungen über Algebra, pp. [illegible].

[illegible] Let this [illegible] then [illegible] interval [illegible] and [illegible] in the [illegible] we can [illegible] [illegible] $x \neq b$ [illegible] less than [illegible] In the same manner we can find [illegible] [illegible] and [illegible] if $f(x)$ is negative in (b, c).

[illegible] multiplicit[illegible] Let [illegible] the number of [illegible] real roots of [illegible] the number of [illegible] the corresponding numbers m'_1, and m'_2 for $g(x)$ satisfy $\quad m'_1 \geq m_1 - 1$

$$m'_2 \geq m_2 - 1$$

[illegible] interval [illegible] between will [illegible] and one of them would [illegible] a double root [illegible] an interval [illegible] with its own multiplicity. Hence $m'_2 \geq m_2$. So we get the

Lemma. [illegible] the number of different real roots of [illegible] is not less than the number of different real roots of [illegible] and the corresponding proposition holds when each root has been counted with its multiplicity.

[illegible] lemma [illegible] the following theorem:

Theorem [illegible] Let [illegible] be a polynomial with real [illegible] let [illegible] $\neq$ [illegible] $\neq 0$ and [illegible] a polynomial with [illegible] Then

$$g(x) = b_0 f(x) + b_1 f'(x) + \dots + b_m f^{(m)}(x)$$

has not [illegible] different real roots than [illegible] and the corresponding proposition holds when each root has been counted with its own multiplicity.

Proof. Without any loss of generality we suppose that [illegible] $= 1$. Hence for $m = 1$ [illegible] and in this case the theorem [illegible] reduced to the lemma. Let $p >$ [illegible] and let the theorem be proved for $m = p$. We [illegible] prove it for $m = p + 1$. If [illegible] is a root of [illegible] then [illegible] $\neq 0$, and [illegible] $=$ [illegible] where [illegible] has real roots only.

[illegible] the theorem holds for $m = \ldots$ the number of the roots of [illegible] is greater than or [illegible] the number of the roots of [illegible]

$$g_1(x) = \sum [illegible]\, c_i\, f^{(i+1)}(x).$$

If we replace in $h(x)$ [illegible] the powers of [illegible] by the correspond ing derivates of $f(x)$, we get therefore

$$f(x) = g_1(x) - a\, g [illegible]$$

From the preceding [illegible] that [illegible] of the roots of [illegible] than the number of the roots of [illegible]

1.5 GRAEFFE'S METHOD

By Graeffe's method [illegible] [3, 1] [illegible] roots [illegible] have to be obtained [illegible]

Let $$b_1 > b_2 > \ldots > b_n \tag{5, 1}$$

be the roots of the polynomial $a_0x^n + a_1x^{n-1} + \ldots + a_n$, then

$$-\frac{a_1}{a_0} = b_1\left(1 + \frac{b_2}{b_1} + \ldots\right) [illegible]$$

$$\frac{a_2}{a_0} = [illegible]$$

[illegible]

$$\ldots$$

[illegible]

$$\frac{-a_n}{a_{n-1}} = [illegible] = b_n(1 + \ldots$$

If [illegible] is very great for [illegible] the [illegible] can be [illegible]. In this case we get the approximation

$$[illegible] \qquad \text{for } i = 1, \ldots, n. \tag{B. 2}$$

In general [illegible] Therefore we have to [illegible] and by repeating this construction we get a sequence of polynomials with the roots [illegible]

[illegible]

Let [illegible] have the roots [illegible] and [illegible] so great that [illegible] can be [illegible] then [illegible]

[illegible] Hence we have to repeat the [illegible] that after further rep[illegible] the squares of the [illegible]

[illegible]

$$[illegible] (x^2 - b_1^2) \ldots (x^2 - b_n^2)$$

$$= f_2(x^2)$$

The coefficients of f_2 [illegible] calculated by the following scheme

(1)	a_0	a_1	a_2	$\ldots$	a_n
	a_0	$-a_1$	a_2		$\pm a_n$
(2)	a_0^2	$-a_1^2$	a_2^2	$\ldots$	$\pm a_n^2$
		$+2a_0a_2$	$-2a_1a_3$		
			$+2a_0a_4$		

As in the first pair of lines corresponding numbers differ only by the sign, it is usual to write only the signs in the first row. The numbers increase very quickly, therefore it is convenient to omit the last figures denoting the decimals very exactly. For this purpose we also use the notation

$$8^{9}456181 \qquad \text{for} \qquad 8{\cdot}456181 \cdot 10^{9}.$$

To extract the roots at the end of the calculation we need logarithms. It is therefore useless to calculate more decimals than the tables of logarithms contain.

Example. $x^3 - 10x^2 + 16x - 2 = 0$

(1)	1	$-1^{1}0$	$1^{1}6$	-2
	+	+	+	+
	1	$-1^{2}0$	$2^{2}56$	-4
		$+0^{2}32$	$-0^{2}4$	
(2)	1	$-6^{1}8$	$2^{2}16$	-4
	+	+	+	+
	1	$-4^{3}624$	$4^{4}6656$	-16
		$+0^{3}432$	$-0^{4}0544$	
(4)	1	$-4^{3}192$	$4^{4}6112$	-16
	+	+	+	+
	1	$-1^{7}75729$	$2^{9}12631$	-256
		$+0^{7}00932$	$+0^{9}00013$	
(8)	1	$-1^{7}74797$	$2^{9}12618$	-256

In the next step the coefficients will become the squares of the preceding coefficients, and in no case the error will have influence on the first 6 figures. Therefore we stop the procedure and calculate now the roots by the help of logarithms.

log. of the coefficients	log. x^8	log $\lvert x \rvert$	x
0	$\bar{7}$·24254	0·90532	8·0412
$\bar{7}$·242 [illegible]	2·[illegible]506	0·20[illegible]	1·8228
[illegible]	1·[illegible]64—8	0·13[illegible] − 1	0·1365
2·40224			
		0·30103	10·0000
		= log 2	

The signs of the roots cannot be determined by Graeffe's method; we have to arrange a special investigation for the signs in every case. In this example the coefficients have alternating signs, hence the coefficients of $f(-x)$ are all positive. The real roots of $f(-x)$ are therefore negative, hence the roots of $f(x)$ are all positive.

For checking we form the elementary symmetric functions of the approximate roots, and we get

$$s_1 = 10, \qquad s_2 = 15·9998 \qquad s_3 = 2$$

for 10 16 2

§ 5. 2) If a real polynomial has complex roots, two of them are always conjugate, and these have therefore the same absolute value. Graeffe's method has therefore to be modified in this case. An example will give valuable hints for necessary modifications.

Example. $x^4 - 11x^3 + 29x^2 - 24x + 2$.

We know from § 1 that the polynomial has two real roots $b_1 = 7·5026$ and $b_4 = 0·09324$ and two complex roots.

The calculation by Graeffe's method is given * on the next page.

If the procedure be repeated, the two first and the two last coefficients will become the squares of the corresponding coefficients of line (8), but the third coefficient will depend also upon the second and the fourth. We cannot expect that further repetition of the procedure will make the third coefficient independent of his neighbours as two roots of the polynomial have an equal absolute value. If b_1 is greater than the absolute value of

* As the sign in the 2d line is always +, we omit them here for abbreviation.

	1	− 11	29	− 24	2
(1)	1	− 121	841	− 576	4
		58	− 528	+ 116	
			+ 4		
(2)	1	− 63	317	− 400	4
	1	− [illegible]	[illegible]	− [illegible]	16
		+ 634	− 57600	+ 2536	
			8		
(4)	1	− 3335	42537	200064	16
	1	− 1^7 11224	1^8 [illegible]	− 4^{10} 3707	256
		+ 851	− 1 41526	0	
			[illegible]		
(8)	1	1^7 [illegible]	[illegible]	− 4^{10} [illegible]	256

the complex roots, then

$$-\frac{a_1}{a_0} = b_1^m \frac{(1 + 2b_2^m + b_3^m)}{b_1^m} \backsim b_1^m \text{ for a suitable } m .$$

A rough mental calculation shows that $b_1^8 \backsim 60$, $b_1^8 \backsim 1.2$.

The same consideration made for $f(\frac{1}{x})$, shows that if $b_4 < b_3$

$-\frac{a_3}{a_4} \backsim b_4^m$ [illegible]. Hence the complex roots are only dependent on the middle coefficients. In order to get the law of dependence we shall generalise the considerations.

Let B and C be two intervals so that every number of C is very small in comparison to the numbers of B and let

$$f(x) = a_0 + a_1 x + \dots + a_n x^n \quad [illegible]$$

have two sets of roots

$b_1, b_2, \ldots b_s$ whose absolute values belong to B and

$c_1, c_2, \ldots$ [illegible] C

Let [illegible] then [illegible] be [illegible] approximately equal to the [illegible] symmetric fundamental function of $b_1 \ldots b$ [illegible] and

[illegible]

where [illegible] is a suitably chosen mean value of the roots of the second set, and therefore a number of C. Let y be a number of B then

[illegible]

$$\backsim \ldots = \frac{1}{a_r} f_1(y).$$

Let [illegible] and [illegible] be one of the roots b, then $f_1(y) \backsim 0$, and we can therefore approximate the 1st set of the roots of $f(x)$ by the roots of $f_1(x)$. But as a common factor of the coefficients has no importance, the roots b are approximated by the roots of

[illegible]

Let [illegible] then [illegible]

has the roots [illegible] $= b_1$ [illegible] $= b$ [illegible]

and [illegible]

the absolute values of b belong to an interval B, the absolute values of [illegible] belong to C and every number of [illegible] is small in comparison to B'. Hence the roots b can be approximated by the roots of [illegible] $+ a_r$, and therefore the roots [illegible] of [illegible] can be approximated by the roots of

$$a_r x^r + a_{r+1} x^{r-1} + \ldots + a_n$$

So the polynomial [illegible] has to be split up in two polynomials: the first is defined by the $r + 1$ upper terms and leads to the upper class of roots,

the second one is defined by the $n + 1$ lower terms, and leads to the lower class of roots.

The two classes may also be divided into sub-classes etc. Finally we get classes

$$b_1, b_2, \ldots, b_{r_1};\quad b_{r_1+1}, \ldots, b_{r_2};\quad \ldots;\quad b_{r_k+1}, \ldots, b_n,$$

each root being small in comparison with the roots of the preceding classes, and to each class corresponds a polynomial which can be cut out from $f(x)$. The ratio of the absolute value of the roots increases when we replace these roots by higher powers of them; therefore we get finally by Graeffe's method k polynomials each of them having only roots with the same absolute value. In the previous example these polynomials are

$$x - 1^{\cdot}10371 \quad [illegible]\ 71\, x^2 - 2^{8}413\, x + 4^{16}8767, \quad 4^{16}8767\, x - [illegible]$$

From these polynomials we get the roots of $f(x)$:

$8 \log|b_1| = 7{\cdot}01286$ $\qquad \log|b_1| = [illegible]$ $\qquad b_1 = [illegible]2$

$8 \log|b_4| = 7{\cdot}7[illegible] - 10$ $\qquad \log|b_4| = [illegible] - 2$ $\qquad b_4 = [illegible]32$

$16 \log|b_2| = \log 4^{16}8767 - \log 1^{\cdot}10371$

$= [illegible]$ $\qquad \log|b_2| = [illegible]$ $\qquad |b_2| = [illegible]$

$\log\cos 8\varphi = \log 2^{8}413 - 8\log|b_2| - \log 2 - \log 1{\cdot}10371 = [illegible] - 10$

$8\varphi = \pm 73°31' + k\ 360°$

$\varphi = \pm 9°11'23'' + k\ 45°$

To finish the calculation we have to fix the signs of the real roots and to determine the integral number k. As the signs of the coefficients are alternating there is no negative root. Hence $b_1 = 7{\cdot}[illegible]$, $b_4 = [illegible]$. These numbers correspond to the results obtained in [illegible] by Horner's method and by Lagrange's method.

As $b_1 + b_2 + b_3 + b_4 = 11$, $2r\cos\varphi = [illegible]$. But as $2r = 2{\cdot}[illegible]$, φ must be a very small angle. Hence $k = 0$. So we get

$$b_2 = 1{\cdot}0367 + i\,0{\cdot}20808$$

$$b_3 = 1{\cdot}0367 - i\,0{\cdot}20808$$

For checking $\log b_1 + \log b_4 + 2 \log r = 0.30104$

for $\log 2 = 0.30103$

$b_1 + b_2 + b_3 + b_4 = 10.1880$

for 11

If we replace in this last verification the value for b_1 by the more exact value obtained from [illegible] $b_1 = -7.2025$ we will get

$b_1 + b_2 + b_3 + b_4 = 10.9991$. The result can be corrected by further iteration. As we see from the results of [illegible] and from the checking given here b_1, b_4 and r are very exact. The correction is therefore expected to concern mainly the angle ϕ whose true value may be a little smaller. As ϕ itself is a small angle this correction will materially affect $\sin \phi$. Hence the imaginary parts of b_2 and b_3 are true up to the second decimal only.

If a polynomial with roots of equal absolute value has a degree > 2, either it has multiple roots or it has non-conjugate roots. The multiple roots will be removed, when we divide by the h.c.f. of the polynomial and its derivative. Non-conjugate roots of equal absolute value can be cleared away by Horner's scheme, for if $|x| = |x'|$ and x' is different from x and $\bar{x}$, then $|x - a| \neq |x' - a|$.

Hence the real and the complex roots of $f(x)$ can be found out by a combination of Graeffe's method and Horner's scheme in every case. The results should be verified and it is possible to minimise the error by the methods given in § 1.

§ 4. Roots of complex polynomials.

Let $\phi(x)$ be a polynomial with complex coefficients,

$$\phi(x) = a_0 + a_1 x + \dots + a_n x^n$$

$$\bar{\phi}(x) = \bar{a}_0 + \bar{a}_1 x + \dots + \bar{a}_n x^n,$$

$(\phi(x), \bar{\phi}(x)) = f_1(x)$, $\quad \phi(x) = f_1(x)\, \phi_1(x)$, $\quad \bar{\phi}(x) = f_1(x)\, \bar{\phi}_1(x)$

$\phi_1(x)\, \bar{\phi}_1(x) = f_2(x)$,

then $f_1(x)$ and $f_2(x)$ are real polynomials. On applying Graeffe's method to these polynomials we get the roots of $\phi(x)$, but out of two conjugate

roots of $f_2(x)$ one is a root of $\phi(x)$, the other is a root of $\bar{\phi}(x)$, and we have therefore to make a verification finally.

Let $|z| \geq \sum_{i=0}^{n} \frac{|a_i|}{|a_n|} = t$, then

$$\frac{1}{|a_n|}|\phi(t)| \geq t^n - \sum_{i=0}^{n-1} \frac{|a_i|}{|a_n|} t^i \geq t^n - t^{n-1}(t-1) = t^{n-1} > 0$$

hence $\phi(t) \neq 0$, and therefore the absolute value of the roots of $\phi(z)$ is $< t$.

Another limit for the roots can be found out by Kakeya's theorem.

The roots of $\phi(z)$ are also roots of the real polynomial $f_1(z) \cdot f_2(z) = f(x)$ where $z = x + a$ and the real number a can be chosen in such a manner that $f(x) = a_n x^n + \dots + a_1 x + a_0$ has positive coefficients only.

For the polynomials with positive coefficients the following theorem holds.

Theorem. Let the coefficients of $f(x) = a_0 + a_1 x + \dots + a_n x^n$ be positive and $0 < p < \frac{a_{k-1}}{a_k} < q$ for $k = 1, \dots, n$, then the roots of $f(x)$ have to satisfy the condition

$$p < |x| < q.$$

Proof. Let $x = qy$, $f(x) = g(y) = \sum b_k y^k$ then $b_k = q^k a_k$.

Hence $b_{k-1}/b_k < 1$. From Kakeya's theorem it follows therefore, for the roots that $|y| < 1$, and $|x| < q$.

The roots of $F(x) = a_0 x^n + \dots + a_{n-1} x + a_n$ are reciprocal to the roots of $f(x)$. As $\frac{a_k}{a_{k-1}} < \frac{1}{p}$ holds, it follows from the first part of the proof that the roots of F have to satisfy

$|x| < \frac{1}{p}$. Hence $|x| = \left|\frac{1}{x}\right| > p$ holds for the roots of $f(x)$.

An interesting connection between the roots of $\phi(z)$ and its derivate $\phi'(z)$ is given by the

The [illegible] theorem. Every convex polygon including all the roots of $\varphi(x)$ contains every root of $\varphi'(x)$.

Proof. Without any loss of generality we can suppose that φ and φ' have no common root. Let γ be an arbitrary root of φ' and $\beta_1, \ldots, \beta_n$ be the roots of φ, then

$$\frac{\varphi'(x)}{\varphi(x)} = \sum \frac{1}{x-\beta_i}, \quad \text{hence } 0 = \frac{\varphi'(\gamma)}{\varphi(\gamma)} = \sum \frac{1}{\gamma-\beta_i} \quad \text{and therefore}$$

$$0 = \sum \overline{\frac{1}{\gamma-\beta_i}} = \sum \frac{\gamma-\beta_i}{|\gamma-\beta_i|^2} = \sum (\gamma-\beta_i)\cdot b_i, \quad \text{where } b_i \text{ is positive.}$$

We consider the geometrical representation of the complex numbers in the plane. $(\gamma-\beta_i)b_i$ are vectors starting from γ and directed to $\beta_1, \ldots, \beta_n$. As the sum is equal to 0, every component of this sum is equal to 0. Let G be an arbitrary straight line passing through γ. The components of $(\gamma-\beta_i)b_i$ orthogonal to G form a sum equal to zero, hence either the components are all equal to zero or there are components with different sign. In the 1st case the points β_i are all situated on G, in the 2nd case, there are roots of φ on both sides of G. In no case there are roots of φ on one side of G only. Let now P be a convex polygon including all the roots of φ. If γ is outside of P we can draw a straight line G not intersecting P through γ. Hence P and therefore all the roots of φ are situated on the same side of G. Hence γ is not a root of φ'.

Let P_0 be the smallest convex polygon including the roots of φ. (The reader may prove that such a polygon exists and is unique.) P_1 the corresponding polygon defined by φ', …, P_i the smallest polygon containing the roots of $\varphi^{(i)}$. The polygons with higher indices are included in the preceding. $\varphi^{(n-1)}$ degenerates to the point $-\frac{1}{n}\frac{a_{n-1}}{a_n} = \frac{1}{n}\sum \alpha_i$. This point is the centre of gravity of the roots of φ and for the same reason it is the centre of gravity of the roots of φ' and of the roots of each derivate.

§ 5. Interpolation

[5/1] Let

$$\beta_1, \ldots, \beta_{n+1} \tag{5, 1}$$

be $n+1$ *different* elements of an arbitrary field K, and let

$$\lambda_1, \ldots, \lambda_{n+1} \tag{5, 2}$$

be $n+1$ arbitrary elements of K.

We want to find out a polynomial $f(x)$ of $K[x]$ such that

$$f(\beta_i) = \lambda_i \text{ for } i = 1, \ldots, n+1 \text{ and degree } f(x) \leq n.$$

Let $f(x) = \sum \alpha_k x^k$. The polynomial has the proposed property if and only if its coefficients satisfy

$$\sum_{k} \alpha_k \beta_i^k = \lambda_i$$

The determinant of this system of $n+1$ linear equations (see Part II [illegible]) is equal to $\pm \prod_{i<j} (\beta_i - \beta_j)$ and is $\neq 0$ as the $n+1$ elements β_i are supposed to be different. Hence the problem has one and only one solution. This solution can be calculated by the methods exposed in Part I, but it is easier to get it from special cases.

Let $f_i(x)$ be the solution of $\lambda_k = 0$, $\lambda_i = 1$, then $f(x) = \sum_{i=1}^{n+1} \lambda_i f_i(x)$

is the solution for arbitrary λ-elements. But $f_i(x) = \frac{g(x)}{(x - \beta_i)\, g'(\beta_i)}$

where $g(x) = \prod_{i=1}^{n+1} (x - \beta_i)$. So we get Lagrange's formula of interpolation

$$f(x) = \sum g(x) \frac{\lambda_i}{(x - \beta_i)\, g'(\beta_i)} \qquad [illegible]$$

By Lagrange's formula the problem of interpolation has been solved in the most complete and general manner, but the formula is not convenient for practical calculation. It is easier to calculate the coefficients of the product representation of $f(x)$ [5/]

$$f(x) = \gamma_0 + \gamma_1 (x - \beta_1) + \gamma_2 (x - \beta_1)(x - \beta_2) + \ldots + \gamma_n (x - \beta_1) \ldots (x - \beta_n) \qquad [illegible]$$

Here is $\gamma_0 = f(\beta_1) = \lambda_1$, $\gamma_1 = \frac{\lambda_2 - \lambda_1}{\beta_2 - \beta_1}$ and we may successively calculate the coefficients γ_i. It is convenient to arrange the calculation in the following manner.

Let $f_k(x)$ be defined by $f_0(x) = f(x)$, and for $k = 1, \ldots, n$

$$f_k(x) = \frac{f_{k-1}(x) - f_{k-1}(\beta_k)}{x - \beta_k},$$

then $f_k(x) = \gamma_k + \gamma_{k+1}(x - \beta_{k+1}) + \ldots + \gamma_n (x - \beta_{k+1}) \ldots (x - \beta_n)$

Hence $f_k(\beta_{k+1}) = \gamma_k$. We have therefore to calculate the values

$$\{k, m\} = f_k(\beta_m) \quad \text{for } k = 0, \dots, n, \qquad k < m \leq n+1$$

by $\{k, m\} = [\{k-1, m\} - \{k-1, k\}] / (\beta_m - \beta_k)$

and $\{0, m\} = \lambda_m$. We calculate the values column-wise in the following scheme

	$\{0, m\}$	$\{1, m\}$	$\{2, m\}$	$\dots$	$\{n, m\}$
$\{k, 1\}$	λ_1				
$\{k, 2\}$	λ_2	$\dfrac{\lambda_2 - \lambda_1}{\beta_2 - \beta_1}$			
$\{k, 3\}$	λ_3	$\dfrac{\lambda_3 - \lambda_1}{\beta_3 - \beta_1}$	$\dfrac{\{1, 3\} - \{1, 2\}}{\beta_3 - \beta_2}$		
$\{k, n+1\}$	λ_{n+1}	$\dfrac{\lambda_{n+1} - \lambda_1}{\beta_{n+1} - \beta_1}$	$\dfrac{\{1, n+1\} - \{1, 2\}}{\beta_{n+1} - \beta_2}$		$\dfrac{\{n-1, n+1\} - \{n-1, n\}}{\beta_{n+1} - \beta_n}$

The first elements of the different columns of this scheme form the set $\gamma_0, \gamma_1, \dots, \gamma_n$ of the coefficients of (4). This scheme is easier for calculation than Lagrange's formula.

[5/8] The reckoning can further be simplified if the elements $\beta_1, \dots, \beta_{n+1}$ are equidistant, i.e. if

$$\beta_{k+1} - \beta_k = \Delta$$

for every k; then

$$\Delta \cdot \{k, m\} = [\{k-1, m\} - \{k-1, k\}] / (m-k)$$

$$= \frac{1}{m-k} \sum_{i=k}^{m-1} \Delta_{k-1, i},$$

where $\Delta_{k-1, i} = \{k-1, i+1\} - \{k-1, i\}$ is the difference of two consecutive elements in the preceding column. So $\Delta \{k, m\}$ is the mean value of the differences of consecutive elements of the rows m to k in the column $k-1$.

We will now transform the scheme for these cases in such a manner that we have to calculate the differences of consecutive elements only and not the mean-values.

For this purpose we introduce the notations usual in the calculus of differences.

Let $\Delta_x \cdot u = x - \beta_1$, then $\Delta_x(u+k-1) = x - \beta_k$.

Let $F(u) = f(x) = f(\Delta_x u + \beta_1) = \gamma_0 + \gamma_1 u\Delta_x + \gamma_2 u(u-1)\Delta_x^2$
$+ \dots + \gamma_n u(u-1)\dots(u-n+1)\Delta_x^n$

and let $\Delta f(x) = f(x+\Delta_x) - f(x) = F(u+1) - F(u)$, then

$$\Delta f(x) = \Delta_x[\gamma_1 + 2\gamma_2 u\Delta_x + \dots + n\gamma_n u(u-1)\dots(u-n+2)\Delta_x^{n-1}]$$

for $(u+1)u(u-1)\dots(u-k+1) - u(u-1)\dots(u-k) = (k+1)u(u-1)\dots(u-k+1)$.

Let $\Delta(\Delta f(x)) = \Delta^2 f(x), \dots, \Delta\Delta^r f(x) = \Delta^{r+1} f(x)$,

then we get by repetition of this procedure

$$\Delta^2 f(x) = \Delta_x^2[2\gamma_2 + 2\cdot 3\gamma_3 u\Delta_x + \dots + n(n-1)\gamma_n u(u-1)\dots(u-n+3)\Delta_x^{n-2}]$$

$$\Delta^n f(x) = \Delta_x^n\, n!\,\gamma_n$$

For abbreviation we shall write $\Delta^r f(\beta_i) = \Delta_i^r$. Then

$$\Delta_i^{r+1} = \Delta_{i+1}^r - \Delta_i^r$$

and $f(\beta_1) = \gamma_0$

$$\Delta_1^1 = \Delta_x \gamma_1$$

$$\Delta_1^k = \Delta_x^k k!\,\gamma_k$$

$$\dots$$

$\Delta_i^n = \Delta_x^n n!\,\gamma_n$ (for $i = 1, 2, \dots$) holds.

So we get *Newton's formula*

$$f(x) = f(\beta_1) + \Delta_1^1 u + \frac{1}{2!}\Delta_1^2 u(u-1) + \dots + \frac{1}{n!}\Delta_1^n u(u-1)\dots(u-n+1)$$

$$= f(\beta_1) + \frac{\Delta_1^1}{\Delta_x}(x-\beta_1) + \frac{1}{2!}\frac{\Delta_1^2}{\Delta_x^2}(x-\beta_1)(x-\beta_2) + \dots + \frac{1}{n!}\frac{\Delta_1^n}{\Delta_x^n}(x-\beta_1)\dots(x-\beta_n)$$

The elements Δ_i^j can be calculated very easily by the following scheme —

$$\begin{array}{ccccc}
\lambda_1 & & & & \\
& \Delta_1^1 & & & \\
\lambda_2 & & \Delta_1^2 & & \\
& \Delta_2^1 & & & \\
\lambda_3 & \cdot & & \Delta_1^3 & \\
\cdot & \cdot & \Delta_{n-2}^2 & & \\
& \Delta_{n-1}^1 & & & \\
\lambda_n & & & &
\end{array}$$

The degrees of $f(x)$, $\Delta f(x)$, ..., $\Delta^n f(x)$ are decreasing and the last one is a constant, so we can use the above scheme also for *extrapolation* to get the value of $f(x)$ for every arbitrary integral value of x, that means for every value $x = a_1 + k\Delta$, where k is an arbitrary integral number.

Example. Let $f(x)$ be of degree 4 and let $f(1)=3$, $f(2)=4$, $f(3)=5$, $f(4)=1$, $f(5)=2$. In order to get $f(6)$ we use the scheme, calculating at first the numbers above the dotted line from the left to the right and then the numbers below the dotted line from the right to the left.

$$\begin{array}{cccccc}
1 & 3 & & & & \\
& & 1 & & & \\
2 & 4 & & 0 & & \\
& & 1 & & -5 & \\
3 & 5 & & -5 & & 15 \\
& & -4 & & 10 & \\
4 & 1 & & 5 & & 15 \\
& & 1 & & 25 & \\
5 & 2 & & 30 & & \\
& & 31 & & & \\
6 & 33 & & & &
\end{array}$$

Hence $f(6)=33$.

PART V

MATRICES. RESULTANTS.

§ 1. Matrices.

In the first part of these lectures matrices have been used to solve [1 1] systems of linear equations and in § 14 and § 15 a few properties of matrices have been discussed. We will now consider the matrices in a more systematic manner.

Let K be an arbitrary field (see Part II, § 2), let 0 be the nullelement, 1 be the unitelement of K, and let

$$a, b, \ldots, f, c, f \text{ with and without indices of any kind} \tag{1, 1}$$

be arbitrary elements of K. A scheme of m rows and n columns*

$$\begin{pmatrix} a_1^1 & \ldots & a_1^n \\ a_m^1 & \ldots & a_m^n \end{pmatrix} = (a_i^k) = A \tag{1, 2}$$

has been called (see Part I, § 6) a *matrix* and a_i^k its *elements*. If $m = n$ this number will be said to be the *degree* of A; the general case can be reduced to the case $m = n$.

Let A, B, C, ... be matrices of degree n and let the elements be denoted by the corresponding small latin type as in (1, 2). The addition of matrices is given by

$$A + B = F, \text{ where} \tag{1, 3}$$

$$a_i^k + b_i^k = f_i^k \qquad \text{for } i = 1, \ldots, n;\ k = 1, \ldots, n.$$

The commutative law $\qquad A + B = B + A \qquad$ (1, 4)

and the associative law $\qquad (A + B) + C = A + (B + C) \qquad$ (1, 5)

hold for the addition of matrices.

Let c be an arbitrary element of the field K, then we define the product

$$c\,A = ((c\,a_i^k)) \tag{1, 6}$$

i. e., we multiply every element of A with c and get a new matrix $c\,A$ of degree n. Then the

distributive laws $\qquad (c + d)\,A = c\,A + d\,A$

$$c\,(A + B) = c\,A + c\,B \tag{1, 7}$$

* Instead of brackets sometimes vertical double bars are put.

hold. For $c=0$, we get

$$0\,A=O, \tag{1, 8}$$

the [illegible] of O [illegible] equal [illegible]

$$A + [illegible] A \tag{1, [illegible]}$$

If we define the *subtraction* by

$$A-B=A+(-1)\,B, \tag{1, 10}$$

the [illegible] from a [illegible] Part [illegible], M in which the [illegible]. In M the [illegible] with an element of K is a [illegible] Part I [illegible].

[illegible] M is a [illegible] K, see Part II [illegible]. M will be proved [illegible] by [illegible] independent matrices E [illegible]. Let $E\,(r,s)$ be defined by

$$e^r_s\,(r,s)=1$$

$$e^i_k\,(r,s)=0 \qquad \text{for } (i,k) \neq (r,s) \tag{1, 11}$$

then

$$\sum_{i,k} a^i_k\,E\,(i,k) = A \tag{1, 12}$$

[illegible] and A is [illegible] and only [illegible] every [illegible] equal to 0. Hence the n^2 matrices are [illegible] independent. They form a [illegible] of M, and M can be [illegible] of rank n^2, see Part I [illegible] 4, Part II [illegible]. By the [illegible] K the [illegible] becomes uniquely defined up to an arbitrary isomorphism.

[1, 2] In Part I [illegible] the [illegible] of matrices of M has been defined by

$$A\,B=D, \quad d^i_k=\sum_j a^i_j b^j_k \tag{1, 13}$$

and [illegible] law $AB\,C = A\,BC$ [illegible] (1, 14)

has [illegible] been proved. From [illegible] and (1, [illegible]) the

[illegible] laws $\quad (A+B)\,C = A\,C + B\,C$

$$C\,(A+B)=C\,A+C\,B \tag{1, 15}$$

follow directly.

M is therefore a [illegible] Part II [illegible]. The reader may prove as an [illegible] that the ring is non-commutative except when $n=1$. The ring has as the unitelement, the matrix

$$E = e^i_k \quad \text{where } e^i_i = 1 \text{ and } e^i_k = 0 \text{ for } i \neq k \tag{1, 16}$$

Then $$E\,A = A\,E = A,\qquad (1,17)$$

and $$O\,A = A\,O = O\qquad (1,18)$$

for every matrix A. [illegible] $B \neq I$ [illegible]

E.g. let A [illegible]

To every matrix A there [illegible] (see Part I § 16

$$\det(AB) = \det A \cdot \det B\qquad (1,19)$$

On the other hand, [illegible] a matrix A for which $\det A = 0$ [illegible] the multiplication remains invariant.*

[illegible]

It is therefore [illegible] elements of K by the [illegible] mentioned that [illegible]

$$\det(cA) = c^n \det A$$

[illegible]

$$cB \cdot C = c\,BC,\qquad (1,21)$$

If $\det A$ [illegible] and let

$$L_j = ab_j\qquad (1,22)$$

be the cofactor of a_j (see Part I p. 23).

* The [illegible]

then $\quad \sum a_j^i b_k^j = 0 = \sum a_i^j b_j^k \quad$ for $\quad i \neq k$

$\sum$ [illegible] $= 1 = \sum$ [illegible] holds [see Part I [illegible] and [illegible]]. Hence

$$AB = E = BA \tag{1, 22}$$

holds. The matrix B the elements of which have been defined by (1, 22) is said to be the inverse of A and will be denoted by A^{-1}.

$$\text{So} \qquad AA^{-1} = E = A^{-1}A \tag{1, 23}$$

There exists therefore an inverse matrix if and only if the determinant is different from 0.

Let $\det A \neq 0$, then from

$$AX = B \; , \qquad YA = B \quad \text{it follows that}$$

$$X = A^{-1}B, \qquad Y = BA^{-1}$$

Hence the matrices with non-vanishing determinant form a set in which the two inverse operations of the multiplication [illegible] and give a unique result. [illegible] the determinant of the [illegible] matrices with non-vanishing determinants may be equal to 0. Of course every matrix [illegible] non-vanishing determinant [illegible] [illegible].

(1/8) We have to consider now [illegible] matrices. To every matrix A there exists a linear transformation of the space of the n vectors [illegible] transforming the vector [illegible]

$$a_1^i x_1 + a_2^i x_2 + \ldots + a_n^i x_n = x'_i \tag{1, 24}$$

$$i = 1, \ldots, n$$

By this transformation the unit vectors (see Part I p. [illegible]) become transformed to the vectors [illegible] defined by the kth column. If we want to write formula (1, 24) as a matrix formula it is convenient to introduce a special notation for matrices which have every element outside the first column equal to 0.

Let

$$\begin{pmatrix} x_1 & 0 & \ldots & 0 \\ x_2 & 0 & & 0 \\ & & & \\ x_n & 0 & & 0 \end{pmatrix} = x \tag{1, 25}$$

then (1,24) becomes

$$A(x) = (x') \qquad (1,24')$$

$$(x) = A^{-1} \text{[illegible]}$$

The formula [illegible] and [illegible] express the transformation by which the basis of the n vectorspace formed by the unit vectors [illegible] becomes transformed to the vectors $\sum a_{ij} e^i$.

Let $x = B y$ [illegible] $= B$ [illegible] and det $B \neq$ [illegible], $\qquad$ (1,26)

then $$AB(y) = B(y')$$

$$B^{-1}AB(y) = (y'). \qquad (1,27)$$

By (1,26) a linear [illegible] correspondence [illegible] the vectors of the vectorspace has been set up [illegible] the unit vectors of the [illegible] there correspond the vectors

$$(B_i) = (b_i^1, \ldots, b_i^n) \qquad (1,28)$$

[illegible] these vectors form a basis and conversely each basis of the vectorspace can be chosen as a set of vectors [illegible] defines the columns of a matrix B with non-vanishing determinant. The vectors [illegible] $B y$ become transformed by (1,27) in the same manner as the unit vectors become transformed by (1,24) [illegible] $\sum$ [illegible]. So (1,27) gives the transformation of the vectorspace if the transformation of an arbitrary basis is given.

The matrix $$B^{-1}A\,B \qquad (1,29)$$

is said* to be the *transform* of A by B. As

$$E^{-1}AE = A,\ B^{-1}[\text{illegible}]B^{-1} = A,\ C^{-1}B^{-1}ABC = (BC)^{-1}A\,BC,$$

hold the transforms of a fixed matrix A form a class (see Part II [illegible]) and from (1,19) it follows that the matrices of this class have the same determinant. The transformations generated by different matrices of the same class are isomorphic, they correspond to the different bases of the vectorspace. If we replace the unit vectors by the basis formed by the vectors $B_1, \ldots$ [illegible] and we give to the vector defined by y the coordinates $x_1, \ldots, x_n$ of $x = By$, the matrix $C = B^{-1}AB$ becomes replaced by A. If [illegible] is given, we may arrange that A will be reduced to a normal form by a suitable choice of B.

* Some authors denote this matrix as the transform of A by B^{-1}.

§ 2. Transformation of A into a normal-form

[2/1] Let A be a matrix with [illegible] K [illegible] be a suitable algebraic [illegible] and [illegible] element of [illegible] the [illegible] vector [illegible] then [illegible] multiplied with an element [illegible] the matrix A is a diagonal [illegible] matrix transforms the n vectorspace in [illegible] multiplied only with an [illegible] transformation means that [illegible] by its sign. In order [illegible] we have to find out n linear independent vectors [illegible] transformed by A into [illegible] the unit vectors in the same [illegible] transforms the vectors [illegible] Hence $B A B^{-1}$ is a diagonal-matrix.

The conditions

$$a_1^k b_1^i + a_2^k b_2^i + \dots + a_n^k b_n^i = \lambda_i b_i^k \qquad (2,\ 1)$$

for $k = 1$ [illegible] i fixed

form a system of [illegible] equations with the matrix

$$A - \lambda_i E \qquad (2,\ 2)$$

Therefore [illegible] exist [illegible] if and only if

$$\det (A - \lambda_i E) = 0.$$

$$A \det [illegible] + \dots - \sum A^i x + \det A \qquad (2,\ 3)$$

is a polynomial χ_A [illegible] field [illegible] supposed to contain the n roots $\lambda_1, \dots, \lambda_n$ of (2,3) and

$$\chi_A(x) = \Pi (\lambda_i - x). \qquad (2,\ 4)$$

[2/2] For a [illegible] of an arbitrary single matrix and its properties we have [illegible] integral functions of a matrix. The powers of a matrix A will be defined by the help of the multiplication in the usual manner

$$A^0 = E,\ A^1 = A,\ A^2 = A\,A,\ \dots,\ A^{i+1} = A^i A \qquad (2,\ 5)$$

and if det A [illegible] [illegible] we [illegible]

$$A^r A^s = A^{r+s} = A^s A^r. \tag{2,6}$$

The powers of A are commutative matrices.

Let $\phi(x) = \sum$ [illegible] by

$\phi(A)$ the matrix

$$\phi(A) = \sum a_r A^r = [illegible] + a_1 A + \dots + a_n A^n \tag{2,7}$$

and ϕ will be said to [illegible] integral [illegible] [illegible] be a polynomial in A [illegible] it follows that

$$\phi(A)\,\psi(A) = \omega(A) = \psi(A)\,\phi(A). \tag{2,8}$$

Hence the integral functions [illegible] ring. We cannot expect that [illegible] degree [illegible] as the ring [illegible] in [illegible] In this ring there are [illegible] hence the commutative ordering cannot contain [illegible] independent matrices. The matrices

$$E,\ A,\ A^2,\ \dots,\ A^{n^2}$$

therefore cannot be independent, and there must be a polynomial [illegible] of degree $\leqq n^2$, with the property that

$$\omega(A) = O.$$

So the matrices may be considered as [illegible] polynomial

λ_k [illegible] is said to be the [illegible] of A [illegible]

λ_i of χ [illegible] there corresponds at least one solution [illegible]

$\neq (0, \dots, 0)$ of (2,1).

Theorem [illegible] If λ_1, [illegible] λ_m are [illegible] a solution of (2,1) corresponding to [illegible] then the vectors $(\beta_1), \dots, (\beta_m)$ are independent.

Proof. Let the theorem not be true [illegible] $\leqq m$ vectors say [illegible] which are dependent but every [illegible]

of them is independent. Then an equation

$$a_1(\beta_1) + \ldots + a_r(\beta_r) = (0) \qquad (2,9)$$

[illegible] by A

we get

$$a_1\lambda_1(\beta_1) + \ldots + a_r\lambda_r(\beta_r) = (0).$$

[illegible] $= (0)$,

[illegible] pendent [illegible]

[illegible]

[illegible] all different A can be

transformed into the diagonal-matrix

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}. \qquad (2,10)$$

[illegible] independent, [illegible] transformed by A. Hence [illegible]

[illegible] does not hold in every case. E.g., Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ then [illegible] $= 1 - x^2$, $\lambda_1 = \lambda_2 = 1$

[illegible] $= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has the rank 1, [illegible] has only one solution $(x_1, x_2) = (1, 0)$.

[illegible] homomorphic to the ring [illegible] see Part II [illegible], although [illegible], from [illegible] $\neq 0$ [illegible] that [illegible] $= 0$. [illegible] may be greater than the degree of the polynomial [illegible] be infinite.

As [illegible]

$\Sigma a_i (B^{-1} A B)^i = B^{-1} \Sigma a_i A^i B$ holds.

Hence from [illegible] follows that [illegible] $= O$ where B is an arbitrary matrix with non-vanishing determinant. In other words

Theorem [illegible]

A similar theorem holds for [illegible]

Theorem [illegible] same charac- [illegible]
teristic polynomial.

Proof. [illegible] $= 1$

$$\det (B^{-1} A B) = \det A \quad \text{holds}$$

$$B^{-1} (A - xE) B = B^{-1} A B - x E \text{; hence}$$

$\det (A - xE) = \det B^{-1} (A - xE) B =$ [illegible]

i.e. $\quad \chi_A (x) = \chi_{B^{-1}AB}(x).$

The two last theorems have a certain [illegible] It will be proved later on that [illegible]
Now we will consider only that case when the roots of the characteristic polynomial are all different.

As [illegible] $(A - \lambda_i E)(\mathfrak{a}_i) = O$. As the n matrices

$A - \lambda_i E$ are commutable [illegible] $= O$ [illegible]

The vector [illegible] the vector space of rank [illegible] rank [illegible]

Hence (see Part I § 14):

$$\chi_A (A) = O. \tag{2,11}$$

Let [illegible] multiple root of the characteristic polynomial of A [illegible] it is a multiple root of the same order [illegible] characteristic polynomial of $B A B^{-1}$ where B is an arbitrary matrix of [illegible] vanishing determinant. [illegible] of (2,1 [illegible] Let [illegible] be the first column of a matrix B [illegible] transformed by [illegible] in the same manner as [illegible]
Therefore

$$B A B^{-1} = \left\| \begin{matrix} \lambda & * & \cdots & * \\ 0 & & A & \end{matrix} \right\| \tag{2,12}$$

In this formula the asterisks * denote certain elements which will not be considered particularly and [illegible] degree $n-1$.

λ [illegible] root of $\chi_{A'}(x)$ of order $r-1$. We therefore can transform A' into

$$B'A'B'^{-1} = \begin{Vmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & A'' & \\ 0 & & & \end{Vmatrix}$$

where A'' is of degree $n-2$, and if

$$B_1 [illegible] = B_2, \qquad (2,13)$$

$$B_2 A B_2^{-1} = \begin{Vmatrix} \lambda & * & * & \cdots & * \\ & \lambda & * & \cdots & * \\ 0 & 0 & & & \\ \vdots & \vdots & & A'' & \\ 0 & 0 & & & \end{Vmatrix}$$

The first [illegible] row [illegible] the second row [illegible] $r-2$ [illegible] χ_A [illegible] the procedure till we get

$$[illegible] \qquad (2,14)$$

A [illegible] not divisible [illegible] B_r are transformed [illegible] are transformed by $B_r A B_r^{-1}$. Hence

$A(\beta_1) = \lambda(\beta_1)$

$A(\beta_2) = \lambda(\beta_2) + c(\beta_1)$

$A(\beta_3) = \lambda(\beta_3) + d_1(\beta_1) + d_2(\beta_2)$

$A(\beta_r) = \lambda(\beta_r) + k_1(\beta_1) + \ldots + k_{r-1}(\beta_{r-1})$.

Therefore

$(A - \lambda E)(\beta_1) = 0$ [illegible]

$(A - \lambda E)(\beta_2) = c(\beta_1)$, $(A - \lambda E)^2 (\beta_2) = (0)$

$(A - \lambda E)(\beta_3) = [illegible]$ [illegible] $= 0$

[illegible] $(A - \lambda E)^{r}(\beta_r) = (0)$,

Hence for every vector (α) of the vectorspace V generated by (β_1) [illegible] (β_r)

$$(A - \lambda E)^r (\alpha) = (0) \qquad (2, 16)$$

[illegible] The vectors [illegible] are columns of a matrix with non-vanishing determinant, hence they are independent. Therefore rank $V = r$.

We want to prove now that every vector (γ), for which

$$(A - \lambda E)^q (\gamma) = (0) \qquad (2, 17)$$

holds belongs to V. If there would be such a vector (γ) outside of V, we can choose it so that $(A - \lambda E)(\gamma)$ is a vector of V [illegible]

$$A(\gamma) = \lambda(\gamma) + h_1(\beta_1) + \ldots + h_r(\beta_r).$$

As (β_1) [illegible] (β_r), (γ) are supposed to be independent there exists a matrix C, of which these vectors form the $(r+1)$ first rows.

Then

[illegible]

and therefore $\chi_A(x) = (\lambda - x)^{r+1} [illegible]$, contrary to the supposition that r is the highest exponent of $(\lambda - x)$ in $\chi_A(x)$. The vectorspace (2, 17) is therefore composed of all vectors which satisfy (2, 16) for any exponent q. As $(A - \lambda E)(\alpha) = (\alpha')$ satisfies (2, 17) if (α) satisfies it, $A(\alpha) = (\alpha') + \lambda(\alpha)$ is a vector of V if (α) belongs to V. By these considerations we get the following theorem.

Theorem. If λ is a root of $\chi_A(x)$ of order r, then the vectors satisfying (2, 17) for any q form a vectorspace V of rank r. Each vector (α) of V satisfies (2, 16) and is transformed by A to a vector of V. If the first r columns of B form an arbitrary chosen basis of V, then (2, 14) holds

[2/5] Let

$$\chi_A(x) = (\lambda_1 - x)^{r_1} \dots (\lambda_m - x)^{r_m}, \qquad (2, 18)$$

where $\lambda_1, \dots, \lambda_m$ are different. To every λ_i there corresponds a vectorspace V_i of rank r_i, so that for every vector (a_i) of V_i, the equation $(A - \lambda_i E)^{r_i}(a_i) = (0)$ holds and (a_i) is transformed by A into a vector of V_i. To prove that the vectorspaces V_i are independent, we have to use the following lemma.

Lemma. The h. c. f. $\varphi(x)$ of the polynomials $\varphi_1(x), \dots, \varphi_m(x)$ of $\Lambda[x]$, can be expressed by $\varphi(x) = \varphi_1(x)\psi_1(x) + \dots + \varphi_m(x)\psi_m(x)$, where $\psi_1, \dots, \psi_m$ are polynomials of $\Lambda[x]$.

Proof. The lemma is true if $m = 1$ and $m = 2$ (see Part II, 4/5); let it be true for $m = r - 1$ and therefore the h. c. f. of $\varphi_1, \dots, \varphi_{r-1}$ be $\omega(x) = \varphi_1(x)\psi_1(x) + \dots + \varphi_{r-1}(x)\psi_{r-1}(x)$. As the h. c. f. φ of $\varphi_1, \dots, \varphi_r$ is the h. c. f. of ω and φ_r, there is $\varphi(x) = \omega(x)\rho(x) + \varphi_r(x)\psi_r(x) = \Sigma\, \varphi_i(x)\, \psi_i(x)$.

The preceding lemma will be applied to the polynomials

$$\varphi_i(x) = \chi_A(x) : (x - \lambda_i)^{r_i}.$$

As these polynomials are relatively prime, there exist m polynomials $\eta_i(x)$ satisfying

$$\eta_1(x) + \dots + \eta_m(x) = 1 \qquad (2, 19)$$

$\eta_i(x)$ is divisible by $\varphi_i(x)$ ($i = 1, \dots, m$) and therefore by $(x - \lambda_k)^{r_k}$ for $k \neq i$.

Hence $\eta_i(A)(a_k) = (0)$ and from (2, 19) it follows that $\eta_i(A)(a_i) = (a_i)$. (2, 20)

If m vectors (a_i) of the different vectorspaces V_i satisfy

$$(a_1) + \dots + (a_m) = 0,$$

we get by multiplication with the matrices $\eta_i(A)$ for every i

$$\eta_i(A)(a_i) = (a_i) = (0).$$

Hence the vectorspaces $V_1, \dots, V_m$ are independent.

If $(a_i^1), \ldots, (a_i^{r_i})$ is a basis of V_i, the $n = r_1 + \ldots + r_m$ vectors

$$(a_1^1), \ldots, (a_1^{r_1}), (a_2^1), \ldots, \ldots, (a_m^{r_m}) \tag{2, 21}$$

form a basis of the space of the n vectors. The matrix C whose columns are the vectors (2.21) has therefore a determinant $\neq 0$. The vectorspaces V_i are invariant for the transformation A. Hence the m vectorspaces generated by

$$e_1, e_2, \ldots, e_{r_1}$$

$$e_{r_1+1}, \ldots, e_{r_1+r_2}$$

$$\vdots$$

$$e_{n-r_m+1}, \ldots, e_n$$

are invariant for $C^{-1} A C$. This matrix has therefore the form

$$C^{-1} A C = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix} \tag{2, 22}$$

where A_i is a matrix of degree r_i and its characteristic polynomial is equal to $(\lambda_i - x)^{r_i}$. Every element situated outside the squares is equal to 0. As it has been proved in [2.4], the equation $(\lambda_i - A)^{r_i}(a_i) = 0$ holds for every vector (a_i) of V_i. As the vectors (2.21) form a basis of the total vectorspace every vector a is the sum of vectors a_i. Hence

$(\lambda_1 - A)^{r_1} \ldots (\lambda_m - A)^{r_m}(a) = 0$ holds for every vector (a).

Hence $\chi_A(A) = O$.

The essence of these considerations is given by the following theorem.

Theorem. Let (2.18) be the characteristic polynomial of A, then there exist m independent vectorspaces V_i of rank r_i $(i = 1, \ldots, m)$. These vectorspaces are invariant for A. If the columns of C are bases

vectors (2, 23) of the vectorspace V, the transformed matrix $C\,A\,C^{-1}$ is of the form (2, 2[illegible]) [illegible] A is a root of its own characteristic polynomial.

[2/0] If the basis of V is replaced by another basis, the matrix A_i of (2, 22) will be transformed according to the matrices A_i, [illegible] remaining unaltered. To put the matrix (2, [illegible]) into its normal form it is therefore only necessary to transform the matrices A_i [illegible]. Hence there is no loss of generality if we suppose that the characteristic polynomial of A is equal to

$$\chi_A(x) = (x-\lambda)^n.$$

In this case to each vector [illegible] there exists [illegible] $\leq$ [illegible] for which

$$(A-\lambda E)^q(\beta) = (0) \qquad (2, 23)$$

[illegible] the smallest non-negative integer [illegible] for which (2, 23) holds, will be said to be the *exponent* of (β),

$$q = \exp(\beta). \qquad (2, 24)$$

Let $0 < p < q$, $c \neq 0$, then

$\exp(c\beta) = q$, $\exp[(A-\lambda E)^p(\beta)] = q-p$, and

[illegible] d_1 [illegible] then [illegible] d_2 [illegible] $= \exp$ [illegible]

[illegible] d_1 [illegible] d_2 [illegible] $d_1 + d_2 \leq \exp$ [illegible]

From these relations it follows that the vectors for which $\exp(\beta) \leq t$ [illegible] form a vectorspace W_t, and that in the sequence of the vectorspaces

$$W_1, W_2, \ldots, W_r = V \qquad (2, 24)$$

every vectorspace is included in the subsequent space; it may also be identical to it. For every $s < r$ the vectors

$$(A-\lambda E)^s(\beta)$$

form a vectorspace which is [illegible] in W [illegible]. The [illegible] of this vectorspace and W_s will be denoted by W_s [illegible]. Hence in the sequence of the vectorspaces

$W_{s,\,r-1}$, $W_{s,\,r-2}$, ..., $W_{s,\,1}$, W_s

every vectorspace is embedded in the subsequent spaces. Let

$$u_{s-1}, \quad u_{s-2}, \quad \ldots, \quad u, \quad u_0$$

be the rank of the vectorspaces

$$W_{1,s-1}, \quad W_{1,s-2}, \quad \ldots, \quad W_{1,1}, \quad W_1$$

then there exists a basis

$$(\beta_1^1), \ldots, (\beta^1_{u_{s-1}}), (\beta^1_{u_{s-1}+1}), \ldots, (\beta^1_{u_1}), (\beta^1_{u_1+1}), \ldots, (\beta^1_{u_0})$$

of W_1 with the property that

$$(\beta_1^1), \ldots, (\beta^1_{u_i})$$

is a basis of $W_{1,i}$ for every index i. [illegible]

Let $u_i \leq l < u_{i-1}$, then there exists a vector (β_l^{i+1}) satisfying

$$(A-\lambda E)^i \, (\beta_l^{i+1}) = (\beta_l^1) \qquad [illegible]$$

and we define (β_l^k), for $1 < k \leq i+1$ by

$$(A-\lambda E)^{i-k+1} \, (\beta_l^{i+1}) = (\beta_l^k) \qquad (2, 26)$$

From (2, 25) it follows that (2, 26) holds for $k = 1$.

Let the vectors (β_l^k) be arranged in a triangular scheme

$$(\beta_1^1) \quad \ldots \quad (\beta^1_{u_{s-1}}) \quad (\beta^1_{u_{s-1}+1}) \quad \ldots \quad (\beta^1_{u_{s-2}}) \quad \ldots \quad (\beta^1_{u_1}) \quad (\beta^1_{u_1+1}) \quad \ldots \quad (\beta^1_{u_0})$$

$$(\beta_1^2), \ldots, (\beta^2_{u_{s-1}}), (\beta^2_{u_{s-1}+1}), \ldots, (\beta^2_{u_{s-2}}) \quad \ldots \quad (\beta^2_{u_1})$$

$$\ldots \qquad (2, 27)$$

$$(\beta_1^s), \ldots, (\beta^s_{u_{s-1}})$$

If we multiply a vector of this scheme from the left side with the matrix $(A-\lambda E)$ we get the vector just above it

$$(A-\lambda E)\,(\beta_l^{k+1}) = (\beta_l^k), \quad \text{hence}$$

$$A(\beta_l^{k+1}) = \lambda(\beta_l^{k+1}) + (\beta_l^k) \text{ holds.}$$

The vectorspace U_l generated by the vectors of an arbitrary column, say the l-th, is therefore invariant under the transformation by A. These vectors

are independent, viz., from

$$(0) = c_1(\beta_1^1) + \ldots + c_s(\beta_s^1) \qquad \text{it follows that}$$

$$(0) = (A - \lambda E)^{[illegible]}(\beta^1) \ldots (\beta^1) \qquad \text{hence } c_s = 0.$$

The [illegible] form together a basis of U and in terms of this basis the transformation [illegible] is given by the matrix

$$\begin{matrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ \ldots & & & & \\ 0 & & \ldots & \lambda & 1 \\ 0 & 0 & \ldots & 0 & \lambda \end{matrix} \tag{2, 28}$$

where the diagonal elements are λ, in the adjacent parallel line the elements are 1 and the other elements are 0. The degree of (2, 28) is equal to $s+1$ where s is given by (2, 2[illegible]). To get the normal form of the transformation A we have to prove that the vectors of (2, 27) are independent and form a basis of V. The vectors of the first row are independent and form a basis of W_1 by definition.

Let $\sum_{i=1}^{s_1} [illegible] \beta^1 = [illegible]$ where β^1 is any vector of W_1 then we get by

multiplication with $A - \lambda E$ that $\sum [illegible] \beta^1 = [illegible]$ but as the vectors (β_i^1) are

independent, $0 = c_1 [illegible]$ holds, and therefore $(\beta^1) = [illegible]$

Hence the vectors of the two first rows are independent. By mathematical induction it follows* that the vectors (2, 27) are independent.

$$(A - \lambda E)(\beta^{[illegible]}) = \sum_{i=1}^{s_{[illegible]}} [illegible], \text{ as it is a vector of } W_{[illegible]}$$

Let $(\beta^{[illegible]}) [illegible] \sum_{i=1}^{s_{[illegible]}} [illegible] \beta [illegible] (\gamma)$ then $(A - \lambda E)(\gamma) = (0)$ hence (γ) is a

vector of W_1 i.e.

$$(\gamma) = \sum_{i=1}^{s_0} k [illegible] \beta [illegible]$$

* The [illegible] by mathematical induction [illegible] as an exercise

β^2) is therefore dependent on the vectors of the two first rows. Hence these vectors form a base of W_2. If [illegible] is a vector of W_{2}, then [illegible] $= (A - \lambda E)$ [illegible] $- \Sigma$ [illegible] belongs to W_1, hence $k_{i > e_1} = 0$.

$(A - \lambda E)$ [illegible] belongs to W_1, hence [illegible]

The vectors $(\beta_1^1), \ldots,$ [illegible] form therefore a base of the vectorspace $W_{s,j}$.

By mathematical induction* it follows that the vectors of the first k rows form a base of W_k, and if we add the vectors [illegible] we get a base of W_{k+1}. So we get that for k [illegible] r, the vectors (2, 25) form a base of the total vectorspace V.

We arrange this base according to the columns, and we transform in such a manner that the vectors of this base become unit vectors. Then A will be transformed into

$$C A C^{-1} = \begin{pmatrix} A^1 & & & & \\ & A^2 & & & \\ & & \ddots & & \\ & & & \overline{A'} & \\ & & & & \ddots \\ & & & & & A^{e_u} \end{pmatrix} \qquad (2, 26)$$

The transformation of the subspace U^i is given by A^i. Hence these matrices have the normal form (2, 24). The degree of the matrix A^i is equal to the length of the corresponding column in (2, 25). [illegible] form a [illegible] sequence. On the other hand if a matrix is given in the normal form (2, 26) (2, 24), the ranks of the vectorspaces W_1, W_2 [illegible] can easily be found out, the rank of W_k being the number of the matrices A^i, of which the degree is not less than k.

Different normal forms therefore belong to different [illegible] of matrices. In the general case we can transform every matrix A [illegible] (2, 26) [illegible] normal form. By these considerations the [illegible] theorem has been proved.

* See p. 58, footnote.

[illegible] Let (2, 18) be the characteristic polynomial of an arbitrary matrix A. Then A can be transformed into a normal form (2, 22), the matrices $A_1, \ldots, A_m$ have [illegible] degrees of the $A^1, \ldots, A^m$ form a non[illegible] sequence and each of them has the form [illegible] remains arbitrary.

§ [illegible] [illegible] POLYNOMIAL

[illegible]

(3, 1) [illegible] then there is [illegible] with the matrix $(A - \lambda E)$.

[illegible] The matrix (2, 22) [illegible] with

$$\begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_r \\ & \gamma_1 & \cdots & \gamma_{r-1} \\ & & \cdots\cdots & \\ & & & \gamma_1 \end{pmatrix} \qquad (3, 1)$$

where $\gamma_1 \neq 0$ and the elements below the diagonal are equal to 0. Hence (2, 22) is commutative [illegible] and there is no other matrix commutative to (2, 22) than the matrices [illegible] every other transformation will not give [illegible] of the vectorspace corresponding to the normal form.

If $a_i = 0$ each of the matrices A^i is of degree 1. Hence [illegible] In this case the vectorspace is identical with W. There are n spaces $W_1, \ldots$, every basis of the vectorspace leads to the normal form. The normal form is therefore commutative to every matrix with determinant $\neq 0$ (of course it is a diagonal matrix with the diagonal elements $= \lambda$). In the general case the investigation of the admissible transformations of the bases [illegible] the investigation of the matrices commutative to the normal form is more complicated and we

will not go into the details here. The more general problem of finding out the matrices commutative to an arbitrary [illegible] matrix A, can easily be reduced to that problem [illegible] if A and B are commutative, C A C^{-1} and C B C^{-1} are commutative too.

The matrix A is a root of the characteristic polynomial $\chi_A(x)$, but it may be that A is a root of a polynomial of lower positive degree. [1/2]

Let

$$\chi_A(x)=(A-\lambda_1 E)^{r_1}(A-\lambda_2 E)^{r_2}\ldots(A-\lambda_m)^{r_m}$$

To $(A-\lambda_1 E)^{r_1}$ there corresponds the minor matrix A_1 in the normal form [illegible] and this minor matrix has the form [illegible] composed by diagonal matrices A^1 [illegible] $A^{\cdot}$. Let s_1 be the degree of A^1, then $1 \leq s_1 \leq r_1$ holds and every vector [illegible] of the vectorspace corresponding to A_1 satisfies the condition

$$(A-\lambda_1 E)^{s_1}(a_1)=(0)$$

Corresponding to this definition we [illegible] the [illegible] s_2 [illegible] s_m, then

$$1 \leq s_k \leq r_k \qquad \text{and}$$

$$(A-\lambda_k E)^{s_k}(a_k)=(0) \qquad \text{hold,}$$

for every vector [illegible] of the vectorspace corresponding to A_k.

Let $$\psi(x)=(x-\lambda_1)^{s_1}\ldots(x-\lambda_m)^{s_m}, \qquad (8,\ 9)$$

then $$\psi(A)\,(\beta)=(0)$$

holds for every vector [illegible] of the [illegible] and therefore [illegible] holds for every vector of the vectorspace.

Hence $$\psi(A)=O, \qquad (8,\ 5)$$

$\psi(x)$ is a factor of $\chi_A(x)$ and is of lower positive degree except in the case when $r_k = s_k$ for every k, which means that the normal form [illegible] of every k shows only one matrix, [illegible]

To prove that A is not a root of a polynomial which is not divisible by $\psi(x)$, we consider the case $\chi_A(x) = (A-\lambda E)$ [illegible] Let $\varphi(x) = (x-\lambda)^{s}$ [illegible] not divisible by $x-\lambda$ and [illegible] smaller than the degree s of A^1. Then $\varphi(A)$ [illegible] where [illegible] holds. Hence the transformation of the vectorspace generated by [illegible] is an automorphism.

and there exists therefore a vector [illegible] such that [illegible] $(B\,y) \ldots > 0$. Hence $\phi(A)$ [illegible] $\neq 0$), therefore $\phi(A) \neq 0$.

Let $A = \begin{Vmatrix} B & \\ & C \end{Vmatrix}$. From the law of multiplication of matrices it follows directly that $A^2 = \begin{Vmatrix} B^2 & \\ & C^2 \end{Vmatrix}$, $A^k = \begin{Vmatrix} B^k & \\ & C^k \end{Vmatrix}$, and

therefore $\omega(A) = \begin{Vmatrix} \omega(B) & \\ & \omega(C) \end{Vmatrix}$ for every polynomial ω.

Hence $\omega(A) = 0$ if and only if $\omega(B) = \omega(C) = 0$.

The corresponding rule holds if A is composed of more than two diagonal matrices. If therefore A has the form (2, 19), $\omega(A) = 0$ holds if and only if $\omega(A_i) = 0$ for $i = 1, \ldots, m$.

Hence $\omega(x)$ must be divisible by every [illegible]. By these considerations the following theorem has been proved:

Theorem [illegible] $\omega(A) = 0$ if and only if $\omega(x)$ is divisible by the polynomial $\psi(x)$, which has been defined by (3, 2).

[3/5] By the characteristic polynomial of A the transformation generated by A becomes uniquely defined only if the roots of the polynomial are all different. If there are equal roots, there are different normal forms and therefore different non-isomorphic transformations corresponding to the same characteristic polynomial.

To get the characteristic polynomial, it is not necessary to put the matrix in the normal-form. As

$$\chi_A(x) = \det(A - xE),$$

the coefficient of x^k in $\chi_A(x)$ is

$$(-1)^k \Sigma A_{n-k,\,ii} \tag{3, 4}$$

where $A_{n-k,\,ii}$ denote the minors of A with $n-k$ rows which are symmetric to the diagonal of A. The minors are invariant to the transformation;

[illegible] these [illegible] and $k = n - 1$

$$\det(A) \quad \text{and} \quad \Sigma a_{j}. \tag{3, 4'}$$

We want [illegible] the [illegible] linear [illegible] complex variable

$$z = \frac{\alpha w + \beta}{\gamma w + \delta} \quad \text{where} \; [illegible]$$

We introduce [illegible] coordinates [illegible]

$$w_{1} = \alpha z_{1} + \beta z_{2}$$

$$w_{2} = \gamma z_{1} + \delta z_{2}$$

[illegible] common factor of [illegible] arbitrary, we can arrange that [illegible] $\beta\gamma$ [illegible] arbitrary

Let $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ [illegible]

Hence $\lambda_{1}\lambda_{2} = 1$, $\lambda_{1} = re^{i\varphi}$, $\lambda_{2} = r^{-1}e^{-i\varphi}$.

(1) $\lambda_{1} \neq \lambda_{2}$, normal form [illegible] $e^{i\varphi}$ 0 [illegible] $e^{-i\varphi}$ (3, 5)

As a factor ± 1 and [illegible] arbitrary, we can choose $1 \leq r$, $0 \leq \varphi < \pi$.

(2) $\lambda_{1} = \lambda_{2} = \pm 1$, [illegible] $+2$, by a suitable choice of the common factor $= \pm 1$ we can arrange that [illegible] $\lambda_{1} = \lambda_{2}$ [illegible]

There are two normal forms

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{3, 6}$$

These transformations are largely discussed in the elements of the theory of functions. The classes of transformations with the normal form (3, 5) are said to be [illegible], especially for $r = 1$ they are called *elliptic*, for $r > 1$ [illegible] they are called [illegible]. The first matrix (3, 6) denotes the identity, the second matrix denotes a [illegible], the infinite point of the complex sphere being the only fixpoint, and the other transformations of this class are called *parabolic* transformations, the only fixpoint being a finite point.

§ 4 Theory of elementary divisors

[4.1] Let S be [illegible]

1. [illegible] $\neq 0$ of S there corresponds an integral number [illegible] (4, 1)

2. If b is a factor of a, then

$$[illegible] \qquad (4, 2)$$

where the [illegible] if and only if [illegible] b are associated.

If [illegible] arbitrary element of S and a_1 is not divisible by a_2 then there exist in S elements [illegible] satisfying

$$[illegible] = r_2, \quad N(a_3) < N(a_2) \qquad (4, 3)$$

[illegible] have been considered in Part II [illegible] and [illegible] of [illegible] these rings are

1. The ring of the integral numbers, if we define N for every [illegible] of the ring are $+1$ and -1.

2. The ring of the integral complex numbers $a + bi$ (see Part II § 11), where a and b are integers. We define N for $a + bi$ by $N(a + bi) = a^2 + b^2$. The units of this ring are $+1$, -1, $+i$, $-i$.

3. The ring $K[x]$ of the polynomials in an arbitrary indefinite x over an arbitrary field K. We define N for every polynomial $f(x)$ which is different from the polynomial [illegible] by $N(f) =$ degree $f(x) + 1$. The unities of this ring are the elements $\neq 0$ of K.

It has been proved in Part II [illegible]: *The factorisation in S is unique and the h. c. f. of two elements a and b can be expressed by*

$$(a, b) = c\,a + d\,b \qquad (4, 4)$$

where c, and d are elements of S.

As we see from the [illegible] [2,6] the h. c. f. of n elements of S can be expressed by

$$(a_1, \ldots, a_n) = c_1 a_1 + c_2 a_2 + \ldots + c_n a_n. \qquad (4, 4)$$

Exercise. [illegible] for which (4, 1), (4, 2) and (4, 3) hold, prove that it is possible to replace the function N by a function N' satisfying the same conditions as N so that $N'(u) = 1$ if and only if u is a unity of S.

We now show that the method of "sweeping out" [illegible] Part I [illegible] on matrices with elements from [illegible] [illegible] owing to the fact that in Part I the elements of the matrices were assumed to be real numbers [illegible] the operation of "division by an element" [illegible]. In this case we have to deal with the elements taken from a ring, the [illegible] has therefore to be replaced by the [illegible]. The single steps of the "sweeping out" [illegible] differ [illegible]. They can be considered as operations with rows and columns and as multiplications from the left side or from the right side [illegible] both interpretations at any step. The matrices we have to consider are the same as in Part I, § 16. [472]

Diagonal matrices D [illegible], where [illegible], and

$$\text{for } i \neq k,\ d_k^i = 0; \qquad (4,5)$$

Elementary matrices [illegible] where

$$e = 1$$

$$e = \lambda \text{ and} \qquad (4,6)$$

$$\text{every other } e_k^i = 0$$

We get

DA by multiplying every row of A [illegible]

AD [illegible] column A [illegible] d,

$E(\lambda)A$ by [illegible] replacing the row [illegible]

$AE(\lambda)$ by column [illegible] replacing the column [illegible]

We consider especially the diagonal matrices

$$U = ((u_k^i)) \qquad (4,7)$$

for which the diagonal elements $u_i^i = u_i$ are units. By replacing each u_i by u_i^{-1} we get the matrix U^{-1} [illegible] are supposed to be units, U^{-1} too is a matrix of the type (4,7).

$$E^{-1}(\lambda) = E(-\lambda)$$

is an elementary matrix. Hence if we get B by multiplying A from the left as well as from the right [illegible] matrices (4,6) and (4,7) we conversely get A by multiplying B by matrices of the same type. B is said to be *congruent* to A, and is denoted as

$$A \backsim B \qquad (4,8)$$

[illegible] $A \backsim A$. Hence the [illegible] each element of this [illegible] being [illegible] to every other (see Part II [illegible]) [illegible] means to be [illegible] A.

[4.3] Let [illegible] be arbitrary but fixed elements of S and $b_1, \ldots, b_m$ be arbitrary elements of S. The elements

$$a_1 b_1 + \ldots + a_m b_m \qquad (4, 9)$$

[illegible] belongs to M, each [illegible] by [illegible]. But as it has been [illegible]. Hence [illegible] element of S [illegible]. There exists therefore [illegible] between the submodules M of S and the [illegible] every M [illegible] generated by the help of only one element.

We have [illegible] submodule generated by minors of A. Let the determinants

$$A_{k1}, \ldots, A_{km_k} \qquad (4, 10)$$

[illegible] of degree k of the matrix A, and let the elements of A be [illegible], belonging to S, and for every fixed k [illegible] by M_k. Let e_k be the highest common factor [illegible] of M_k [illegible] elements of [illegible] are minors of A of degree k [illegible]. Hence

$$A_{k\,i} = a'_1 A_{k-1, i_1} + \ldots + a'_k A_{k-1, i_k}$$

[illegible] M_{k-1}. Hence in the sequence of modules

$$M_1, \ldots, M_n$$

[illegible] of the preceding and e_k is therefore divisible by e_{k-1}. Hence the [illegible] can be represented by the *elementary divisors* e_k in the following manner

$$\delta_1 = e_1$$

$$\delta_2 = e_1 e_2$$

$$\ldots$$

$$\delta_n = e_1 e_2 \ldots e_n \qquad (4, 11)$$

$$\det A = \delta_n = e_1 e_2 \ldots e_n.$$

Theorem. Congruent matrices have the same elementary divisors.

Proof. Any multiplication from the left or from the right side with a matrix U [illegible] that the rows or the columns [illegible] by [illegible] only [illegible] the row addition $A \longrightarrow E_{r,s}(\lambda)A$, those minors of A in which the r^{th} row is [illegible] not [illegible] in which the r^{th} row as well as the s^{th} row occur. Let A_1 be a minor in which the [illegible] and A_2 [illegible] we get by replacing in A_1 the [illegible] of A [illegible] then A_{1} will be transformed into [illegible] M_1 [illegible] the [illegible]

$$E_{r,s}(\lambda)A \longrightarrow E_{r,s}(-\lambda)(E_{r,s}(\lambda)A) = A.$$

M will be transformed [illegible] M_1 [illegible] M_2 [illegible] M_1 [illegible] M_2, it follows that $M_1 = M_2$. Hence [illegible] row addition M [illegible] and the same holds [illegible] for any column addition. [illegible] the theorem holds. Especially [illegible] of the elementary divisors of A [illegible], and [illegible] the same for every matrix of the same class.

By the [illegible] with matrices [illegible] and U the following operations [illegible] [4/4]

(1) [illegible]

(α)	(α)	$(\alpha) - ((\alpha) + (\beta)) = (-\beta)$	$(-\beta)$	(β)
(β)	$(\alpha) + (\beta)$	$(\alpha) + (\beta)$		(α)	(α)

(2) Sweeping out [illegible] by [illegible] (or row additions if [illegible] of the [illegible] the [illegible] of them

$$a, ab_2, \ldots, ab_n \longrightarrow a, 0, \ldots, 0$$

(3) If none of the elements of a row $a_1, \ldots, a_n$ is [illegible] of them we can alter the row by column additions in such a manner that an element will appear with the property

$$N(b) < N(a_i) \quad \text{for every } i.$$

Proof. Let a_1 be an element for which $N(a_1) \leq N(a_i)$ [illegible] and a_2 not be divisible by a_1; then it follows from (4, [illegible]) that there are elements c and b for which

$$a_2 + ca_1 = b, \quad N(b) < N(a_1) \leq N(a_i) \tag{4, 12}$$

holds.

By column addition we can replace e_2 by b.

4. If e_1 is the h.c.f. of the elements of a row and also the h.c.f. of the elements of [illegible] of all the elements of the matrix, we can arrange by row and column additions that an element b will appear in the matrix for which $N(b) < N(e_1)$ holds.

Proof. From [illegible] and [illegible] follows that we can suppose e_1 to be the first element of the first row and the first row and the first column to be swept out. As by hypothesis e_1 is not the h.c.f. of the elements of the matrix which are not [illegible], there exists an element e_2 not divisible by e_1. Let [illegible] and h [illegible] (2). By interchanging the rows and columns e_2 becomes the 2nd element of the 2nd row. Starting from this position we make row and column additions in the following manner:

$$\begin{matrix} e_1 & 0 \\ 0 & e_2 \end{matrix} \longrightarrow \begin{matrix} e_1 & xe_2 \\ 0 & e_2 \end{matrix} \longrightarrow \begin{matrix} e_1 & b \\ 0 & e_2 \end{matrix}.$$

With the help of these operations the "sweep out" can easily be done. Let e_1 be an element of the matrix A for which that function N has a minimum value. If e_1 is different from the h.c.f. e_1 of the elements of A, then we can alter A in such a manner that there appears an element b for which $N(b) < N(e_1)$ holds and $e_1 \nmid b$. This procedure can be repeated. But the function N takes integral positive values only, hence after a finite number of steps the procedure must stop, and that is only possible if an element of the matrix becomes equal to e_1.

By 1) e_1 will be placed on the left upper corner of the matrix, and by the help of 2, the first row and the first column will be swept out. So we get

$$A \backsim \begin{pmatrix} e_1 & \\ & A_1 \end{pmatrix}$$

Each element of A_1 is divisible by e_1, hence the h.c.f. [illegible] is equal to $e_1 e_2$. By repetition of the procedure we get therefore

$$A \backsim \begin{pmatrix} e_1 & & \\ & e_1 e_2 & \\ & & A_2 \end{pmatrix} \backsim \cdots \backsim \begin{pmatrix} e_1 & & & \\ & e_1 e_2 & & \\ & & \ddots & \\ & & & e_1 e_2 \cdots e_n \end{pmatrix}$$

The h. c. f. of the minors of degree k is [illegible]

[illegible]

Hence [illegible] (4, 13)

[illegible] and

$$A \backsim \begin{pmatrix} e_1 & & \\ & e_2 & \\ & & \ddots \\ & & & e_n \end{pmatrix} \qquad (4,\ 14)$$

e_k divisible by e_{k-1}. (4, 15)

If on the other [illegible] (4, 14) [illegible] of A have the [illegible] we get the following theorem.

Theorem. Every matrix A with [illegible] from S [illegible] to a diagonal matrix (4, 14) where e_k [illegible] These elements e_k are [illegible] elementary [illegible] of A. [illegible] there are [illegible] 1, 1 correspondence between the [illegible] and the sets of [illegible] and every set of [illegible] e_k satisfying (4, 15) is an admissible set.

From the preceding theorem [illegible] A is [illegible] (4, 8) congruent to a diagonal matrix [illegible] the [illegible] e_k are [illegible] Hence A is a product of matrices of type (4, 6) and (4, 7) [illegible] determinant is a unity. On the other hand the determinant of a product of matrices (4, 6) and (4, 7) is a unity. The matrices A whose determinants [illegible] are therefore identical with the products of matrices (4, 6) and (4, 7) and as to each matrix (4, 6) and (4, 7) there exists an inverse matrix of the same type, every matrix whose elements belong to S and whose determinant is a unity of S has an inverse of the same type. [illegible] the following theorem holds.

Theorem. $A \underset{\sim}{\backsim} B$ if and only if

$$A = C_1^{-1}\ B\ C_2, \qquad (4,\ 16)$$

where C_1, C_2 are matrices with elements from S, their determinants being unities of S.

[illegible] transformation of a vector space by the matrix A [illegible] given by

$$[illegible] \; D \; [illegible]$$

[illegible]

$$C_1 \, (x) = (y), \quad (x) = C_1^{-1} (y),$$

$$[illegible]$$

then the modal [illegible] by the vectors [illegible] with the modal formed by the vectors [illegible], the modal formed by the vectors [illegible] with the modal formed by the vectors [illegible]. D is supposed to be a diagonal matrix with [illegible] elements, and

$$(y) = D \, (x) \tag{4, 17}$$

holds.

[illegible]. Let [illegible] independent vectors in the Euclidean space, and [illegible] integral numbers. If the vectors [illegible] [illegible] the [illegible] triplet [illegible] [illegible] that [illegible] is the ring of the integral numbers [illegible]

[4/6] We shall now apply the theory of [illegible] with elements from a ring [illegible] to the theory of the [illegible] with elements from a field [illegible]. The properties [illegible] invariants for transformation and the normal form [illegible] will then appear in a new light.

[illegible] the coefficients of a fixed matrix A and [illegible] The polynomials [illegible] form [illegible] satisfying the properties [illegible] and [illegible] The elements of [illegible] are the [illegible] If [illegible] matrix with elements from [illegible] and det C ≠ 0, then

$$C^{-1} A \, C \backsim A, \text{ and}$$

$$C^{-1} A \, C - xE = C^{-1} (A - x E) \, C \backsim A - x E \tag{4, 18}$$

In order to find [illegible] of A − xE, we can therefore without any loss of generality [illegible] that A has the normal form (2, [illegible]) which consists of a diagonal system of [illegible] matrices [illegible]. Every [illegible] corresponds [illegible] of the characteristic polynomial and it has the normal form [illegible]. This means that [illegible] consists of a diagonal system of submatrices, each of them being of the type [illegible], each element of the diagonal is equal to [illegible], the elements just above the diagonal are equal to 1 and all other elements are equal to 0. The reduction of $A - xE$ to a congruent matrix of the type (4, [illegible]) will be effectuated by the help of the

following congruences

[illegible] (4, 19)

and, for $(a, b) = au + bv = 1$,

[illegible] (4, 20)

We [illegible] where [illegible] The [illegible] of the [illegible] altered by these row and [illegible] We start with the two first rows and columns and change them to $\begin{pmatrix} 1 & 0 \\ 0 & (\lambda_i - x) \end{pmatrix}$ the other [illegible] being unaltered [illegible] of this [illegible] the different submatrices [illegible] We [illegible] congruent diagonal systems with the diagonal [illegible]

$$1, \ldots, 1, (\lambda_1 - x)^{e_1}$$

$$1, \ldots, 1, (\lambda_2 - x)^{e'_1}$$

(4, 21)

$$1, \ldots, 1, (\lambda_i - x)^{e^r_i}$$

where [illegible] (4, 22)

[illegible] has the same [illegible] diagonal matrix with the elements (4, 21), and [illegible] by interchanging the rows and the columns as follows

$$1, \ldots, 1, (\lambda_1 - x)^{e_1}, \ldots, (\lambda_i - x)^{e'_i}, (\lambda_i - x)^{e^r_i}. \qquad (4, 23)$$

From (4, 22) [illegible] that (4, 23) is the normal form [illegible] As the alterations effected in the rows and columns of [illegible] after the [illegible] matrices of A we can reduce [illegible] independently to normal form (4, 2) [illegible] diagonal matrix congruent to A [illegible] the diagonal elements being powers of [illegible] with non-negative exponents. By interchanging the rows and columns and applying the congruence (4, 20) we get a diagonal matrix

$$1, \ldots, 1, \psi^s(x), \ldots, \psi^1(x), \psi(x). \qquad (4, 24)$$

where [illegible] is the same [illegible] has been defined by [illegible] every element [illegible] the preceding elements. Hence [illegible] is the normal form [illegible] of A [illegible]. If [illegible] we can find out the roots [illegible] and the exponents [illegible] and by them the normal form [illegible] defined by these considerations we get the connection between the classes of matrices A formed with the elements of a field K equivalent [illegible] transformations and the classes of matrices formed with elements of the ring K [illegible] and congruent to [illegible]. There is an (1, 1) correspondence between the classes. The normal form [illegible] is directly given by the normal form [illegible] of A and conversely we get [illegible] from [illegible] the [illegible] into products of powers of linear factors.

[illegible]

(6/1) Let K be a sub-field of Λ and (Λ : K) = 2.

Let [illegible] be an [illegible] Λ [illegible] K [illegible] Λ [illegible] the [illegible] K and [illegible] Λ [illegible] K [illegible] polynomial [illegible] K [illegible] of degree [illegible]. Every [illegible] automorphism [illegible] elements of K are not altered [illegible] Hence the [illegible] Λ over K [illegible] transformation [illegible]. In the first case every element of Λ [illegible] the automorphism [illegible] In [illegible] every element [illegible] K. The transformation

$$a + b \longrightarrow a + b \qquad (6, 1)$$

[illegible] automorphism of Λ [illegible] the sum the product of two elements [illegible] transformed into the sum the product of the corresponding [illegible] automorphism of Λ different from the identity by which the elements of K are not altered. Hence the transformation [illegible] is independent of the choice of [illegible] elements of Λ corresponding to [illegible] are said to be conjugate. The conjugate of [illegible] is denoted by

[illegible]

$$\bar{\alpha} = \beta \qquad (6, 2)$$

For the purpose of [illegible] it is [illegible] to consider [illegible] simultaneously two different cases.

(1) [illegible] K [illegible] are [illegible]

(2) [illegible] K [illegible] elements are [illegible]

For both cases we make the following [illegible]

Let α, β be an arbitrary pair of elements of Λ [illegible] different from $(0, 0)$; then there exists an element γ [illegible] such that

$$\alpha\bar{\alpha} + \beta\bar{\beta} = \gamma\bar{\gamma}. \qquad (5, 8)$$

This condition is satisfied e.g.

1. if $K = \Lambda$ is the field of the real numbers,

2. if K is the field of the real numbers, Λ the field of the complex numbers.

The theory has [illegible] general characteristic [illegible] the field of the [illegible] numbers and K be the field [illegible] [illegible] there exist an element ϕ satisfying

$$[illegible] + \dots + \alpha_n\bar{\alpha}_n = \beta\bar{\beta}. \qquad (5, 9)$$

As [illegible] belongs to K [illegible] it follows that [illegible] factors, and therefore both should be equal to 0 [illegible] the equations (5, 9) and [illegible] are therefore different from 0 [illegible] any pair of elements of Λ [illegible] any pair of elements of K.

$$[illegible] \qquad (5, 10)$$

especially [illegible]

As the prime field of Λ [illegible] K [illegible] characteristic of Λ is different from 2 and [illegible] The following theorem will be [illegible]

Theorem. Let [illegible] then [illegible]

conditions

$$\sum_i u_i^j \bar{u}_i^l = 0 \qquad \text{for } j \neq l \tag{5, 5}$$

$$\sum_i u_i^j \bar{u}_i^j = 1 \tag{5, 6}$$

and

$$A_k u^j = c_k^j u^j \qquad \text{for } k = 1, \ldots, m. \tag{5, 7}$$

[illegible] there exists a non-trivial [illegible] holds, for [illegible] We can continue this procedure till we get [illegible]

$$(v_1^j, \ldots, v_n^j) \neq (0, \ldots, 0), \ \sum v_i^j \bar{v}_i^l = 0, \ \text{for } j \neq l$$

[illegible] satisfies the conditions (5, 5), (5, 6) and (5, 7).

[illegible] K [illegible] characteristic different from 0 [illegible] holds.

[5/8] [illegible] 5, 7) can be considered as a property of [illegible] of this kind in a simple form it is helpful to use the following notations.

Let A [illegible] be a matrix with elements from [illegible] then

$\dot{A} = (\dot{a}_{ij})$ [illegible] for $\dot{a}_{ij} = a_{ji}$), is the *transposed* of A, (5, 8)

$\bar{A} = (\bar{a}_{ij})$ [illegible] is the conjugate of A (5, 9)

$\overset{\dagger}{A}$ [illegible] (5, 10)

From these formulae it follows that

$$(AB)^{\bullet} = \dot{B}\dot{A}, \qquad \overline{AB} = \bar{A}\bar{B}, \qquad (AB)^{\dagger} = B^{\dagger} A^{\dagger}$$

and for det $A \neq 0$,

$$(\dot{A})^{-1} = (A^{-1})^{\bullet}, \qquad (\bar{A})^{-1} = \overline{A^{-1}}, \qquad (A^{\dagger})^{-1} = (A^{-1})^{\dagger} \tag{5, 11}$$

The equations [illegible], and [illegible] are therefore equivalent with

$$((u^i_j))^{\dagger} = ((u^i_j))^{-1}. \tag{5, 12}$$

A matrix satisfying this condition [illegible]. As the transpose of an [illegible] matrix is also [illegible] unitary matrix [illegible] also satisfy

$$\sum_i u^i_i \bar{u}^i_j = 0, \qquad \text{for } i \neq j \tag{5, 6'}$$

$$\sum_i u^i_j \bar{u}^i_j = 1. \tag{5, 6'}$$

As det $((u^i_j))$ = det $((u^i_j))'$ = det $((u^i_j [illegible] 1 [illegible]$: det $((\bar{u}^i_j))$ holds, hence

$$\det((u^i_j)) = \pm 1. \tag{5, 13}$$

A unitary matrix remains unitary after an arbitrary permutation [illegible] the rows and of the columns. [illegible] product of unitary matrices [illegible] unitary.

If

$$U = \begin{pmatrix} \pm 1 & \\ & \boxed{U'} \end{pmatrix} \tag{5, 14}$$

and one of the matrices U, U' [illegible] unitary, the other [illegible] unitary [illegible]. If the elements [illegible] unitary matrix are elements [illegible] then [illegible] matrix is said to be [illegible] relation used [illegible] theorem [illegible] in the following [illegible].

Theorem [illegible] To every vector [illegible] there [illegible] the property that [illegible] only.

If the elements of a matrix H [illegible] polynomial* belong to Λ and

$$H = H^{\dagger} \tag{5, 15}$$

H is said to be an *Hermitian matrix*.

Let H be an Hermitian matrix [illegible] a unitary matrix [illegible] then H_1 [illegible] has the same characteristic polynomial as H and $H_1^{\dagger}$ [illegible] $H^{\dagger}$ [illegible] H_1. Hence H_1 is Hermitian. Let λ be an [illegible] and [illegible] $\neq 0$

* This condition is [illegible] is satisfied [illegible] closed field, e.g. the field of the complex numbers.

a vector for which $(H-\lambda I)(t)=0$ holds. Such a vector (t) must exist [illegible] 2. Let U be a unitary matrix whose first column differs from (t) only by a factor $\neq 0$. By $H_1 = U^{-1}HU$ the first unit vector is transformed in the same manner as (t) is transformed by H, hence it is multiplied only by λ. The first column of H_1 has therefore the elements $\lambda, 0, \ldots, 0$ and as H_1 is Hermitian,

$$H_1 = \begin{pmatrix} \lambda & \\ & H' \end{pmatrix}$$

Hence λ is an element of K. H' is an Hermitian matrix of degree $n-1$. We can transform it by a unitary matrix in the same manner as H has been transformed:

$$U'^{-1}H'U' = \begin{pmatrix} \lambda' & \\ & H'' \end{pmatrix}. \text{ As the matrix } U_1 = \begin{pmatrix} 1 & \\ & U' \end{pmatrix} \text{ is}$$

also unitary, $$U_1^{-1}H_1U_1 = \begin{pmatrix} \lambda & & \\ & \lambda' & \\ & & H'' \end{pmatrix}$$ is an Hermitian matrix.

After n steps we get H transformed into a diagonal matrix by a matrix which is the product of unitary matrices and is therefore unitary. So we get the following important theorem.

Theorem. An Hermitian matrix H can be transformed by a unitary matrix into a diagonal matrix and the roots of $\chi_H(x)$ belong to K.

Let K be the field of the real numbers, Λ be the field of the complex numbers; then it follows from the theorem that the roots of the characteristic polynomial are real. This holds especially if the matrix is a symmetric real matrix. Hence we can consider any matrix with real elements for which $a_{ij} = a_{ji}$ holds as Hermitian, where $K = \Lambda$ is equal to the field of the real numbers; hence the roots of the characteristic polynomial belong to Λ. Hence we can apply the preceding theorem on this case, the unitary matrices becoming now orthogonal matrices and we get the

Corollary. If in a matrix A with real numbers as elements $a_k^i = a_i^k$ holds, then $\chi_A(x)$ has real roots only and A can be transformed by an orthogonal matrix with real coefficients into a diagonal matrix.

The theory of matrices will now be applied to linear and quadratic [6/4]
forms. We introduce two sets of indeterminates

$$x_1, x_2, \ldots, x_n, \quad y_1, y_2, \ldots$$

$$\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n, \quad \bar{y}_1, \bar{y}_2, \ldots \tag{5, 16}$$

Corresponding elements of the two lines will be said to be *conjugate*. They are supposed to be different if $[K : \Lambda] = 2$, and to be identical if $K = \Lambda$. The automorphism of K defined in [5.1] by which every element of K is replaced by its conjugate will now be extended to an automorphism of the ring

$$\Sigma = K[x_1, \ldots, \bar{x}_1, \ldots] \tag{5, 17}$$

of the polynomials in the indeterminates (5, 16) with coefficients from K. This extension will be made so that every indeterminate (5, 16) becomes replaced by its conjugate. So we have simply to replace in every polynomial each coefficient and each indeterminate by its conjugate. In the case $K = \Lambda$ the automorphism is the identity. In every case there belongs to every element z of Σ a uniquely defined conjugate element $\bar{z}$, and $\bar{\bar{z}} = z$ holds.

In [1, 4], a vector has been represented (see (1, 25)) by a matrix, the first column of which is formed by the coordinates, the elements of the other columns being 0. We will now apply the notations of conjugate matrix and transposed matrix to these special matrices, so that

$$(x) = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ x_n & 0 & \cdots & 0 \end{pmatrix}, \qquad (\bar{x}) = \begin{pmatrix} \bar{x}_1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \bar{x}_n & 0 & \cdots & 0 \end{pmatrix}$$

$$(x)' = \begin{pmatrix} x_1 & \cdots & x_n \\ 0 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{pmatrix}, \qquad (\bar{x})' = \begin{pmatrix} \bar{x}_1 & \cdots & \bar{x}_n \\ 0 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{pmatrix} \tag{5, 18}$$

Let $A = (a_k^i)$, then $A(y)$ is a matrix with the first element

$$\sum a_k^i x_i y_k \tag{5, 19}$$

all other elements being equal to 0. To every matrix A there corresponds

a bilinear form [illegible] and conversely. Let

$$(x) = B(x'), \quad y = C(y'), \quad \text{and}$$

$$B^* AC = A' = ((a'^i_k)), \tag{5, 20}$$

then $\qquad (x')^* B^* AC(y') = (x)^* A(y).$

Hence $\qquad \sum a^i_k x_i y_k = \sum a'^i_k x'_i y'_k. \qquad (5, 20')$

The formulae (5, 20) and (5, 20') give the transformation of bilinear forms.

We will now consider the case where A is an Hermitian matrix, and where x and y are conjugate; then B and C are also conjugate, [illegible] becomes an *Hermitian form*

$$\Sigma a^i_k x_i \bar{x}_k, \qquad \text{where} \qquad a^i_k = \bar{a}^k_i, \tag{5, 21}$$

and $\sum a^i_k x_i \bar{x}_k = \sum a'^i_k x'_i \bar{x}'_k$, where $((a'^i_k)) = A' = \bar{C}^* AC \qquad (5, 22)$

A' is therefore an Hermitian matrix too. In the special case where $K = \Lambda$, $a^i_k = a^k_i$ and $\bar{x}_k = x_k$, the bilinear form becomes quadratic and the Hermitian matrix becomes a symmetric one:

$$\Sigma a^i_k x_i x_k = \Sigma a^i_i x_i^2 + 2 \sum_{i<k} a^i_k x_i x_k. \tag{5, 23}$$

If the characteristic of K is different from 2, every quadratic form in $x_1, \ldots, x_n$ can be represented in the form (5, 23). We will omit the case of characteristic 2 which needs a special treatment. Hence there is an 1,1 correspondence between the quadratic forms and the symmetric matrices. The transformation is done by

$$((a'^i_k)) = A' = C^* AC, \qquad x_i = C x', \quad \sum a^i_k x_i x_k = \sum a'^i_k x'_i x'_k. \tag{5, 24}$$

The formulae for the transformation of quadratic forms reminds us of the transformation of matrices [see 3, 2 [illegible]], but the notion of inverse matrix has been replaced here by the notion of transposed matrix, and it is not necessary that $\det C \neq 0$. On applying the theorem of [5, 5] and its corollary to (4, 22) and (5, 24) we get the following fundamental theorem:

Theorem. Every Hermitian form can be transformed by a unitary transformation into the normal form

$$\Sigma a_i x_i \bar{x}_i. \tag{5, 25}$$

and every quadratic form with coefficients from K can be transformed by an orthogonal transformation with coefficients from K into the normal form

$$\Sigma a_i x_i^2 \tag{5, 26}$$

a_i being elements of K in both the cases.

Let K = A be the field of the real numbers, then every quadratic form can be transformed into [6,5]

$$Q(x) = a_1 x_1^2 + \dots + a_r x_r^2 \quad , \tag{5, 27}$$

where the coefficients are different from zero, by the transformation with an orthogonal matrix. As the determinant of an orthogonal transformation is equal to ± 1, the rank of the matrix A will not be altered by the transformation and is therefore equal to r. If we replace x by $-x$ the formula (5, 27) will not be altered, but the sign of the orthogonal transformation becomes changed. Hence we can transform an arbitrary quadratic form into a normal form $\Sigma a_i x_i^2$ by an orthogonal transformation with determinant $+1$. Such a transformation means in the geometry of the space of n dimensions a rotation through the origin. This transformation is well known in the analytic geometry of conics and of quadrics as the *transformation to the principal axes.*

Let $a_i = \pm b_i^2$, and $b_i x_i = y_i$, then (5, 27) is transformed to a sum of squares with certain signs $\pm$. After a permutation of the indices and replacing y by x we get therefore

$$q(x) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_r^2 . \tag{5, 28}$$

Hence every quadratic form can be transformed to (5, 28) by a non-degenerated linear transformation with real coefficients. The integer r is the rank of matrix of the quadratic form and therefore invariant. We will prove that p is an invariant too.

Theorem. Every quadratic form with real coefficients can be transformed into one and only one normal form (5, 28) by a non-degenerated linear transformation with real coefficients.

Proof. As it has been shown above, the transformation into the normal form is always possible and r is an invariant. We have therefore only to prove that a transformation of $q(x)$ into

$$q(x) = x_1^2 + \dots + x_q^2 - x_{q+1}^2 - \dots - x_r^2$$

by linear non degenerated substitution is possible only if $p=q$. Let $p \neq q$, say $p > q$ without any loss of generality.

$$x_i = b_1^i z_1 + \dots + b_r^i z_r$$

$$z_i = c_1^i x_1 + \dots + c_r^i x_r \qquad i=1, \dots, r$$

$q(x) = q_1(z)$ for corresponding systems $(x_1, \dots, x_r)$ and $(z_1, \dots, z_r)$. The $q + r - p < r$ linear homogeneous equations

$$c_1^k x_1 + \dots + c_r^k x_r = 0 \qquad k=1, \dots, q$$

$$x_i = 0 \qquad i = p+1, \dots, r$$

have a solution $(\xi_1, \dots, \xi_p, 0, \dots, 0)$ different from $(0, \dots, 0)$ as the rank of the matrix of this system of equations is $\leq q + r - p < r$. The corresponding values of $z_1, \dots, z_r$ are $(0, \dots, 0, \zeta_{q+1}, \dots, \zeta_r)$ and different from $(0, \dots, 0)$. Hence $q(\xi) > 0$, $q_1(\zeta) < 0$ in contradiction to $q(\xi) = q_1(\zeta)$.

A quadratic form is said to be *positive definite* if $n = r = p$, it is *negative definite* if $n = r$, $p = 0$, it is *semi-definite* if $n > r$, $p = r$ or 0, and it is *indefinite* if $r > p > 0$.

[5/6] Finally we will give a geometrical interpretation of the last results without going into the details.

In the projective $(n-1)$-dimensional space a quadric becomes represented by

$$\rho \Sigma a_k^i x_i x_k = 0,$$

where $\rho \neq 0$ is an arbitrary factor.

Hence the quadric has one and only one normal form

$$\rho q_1(x) = 0,$$

and the sign of ρ can be fixed in such a manner that $q(x)$ has not fewer positive than negative terms. We get therefore the different types of quadrics in the projective $(n-1)$-dimensional space given by the different normal-forms

$$q(x) = 0 \quad \text{for} \quad r = 1, \dots, n \quad r/2 \leq p \leq r \tag{5, 29}$$

Especially the quadrics without any real point are those for which $p = r = n$. The normal form (5, 29) has the property that every fundamental point of the coordinate system i.e. every point, for which all the coordinates are equal to 0 except x_i, is polar to the opposite hyperplane $x_i = 0$.

In the affine n-dimensional space the quadrics are given by

$$\Sigma a_{ik} x_i x_k + \Sigma b_i x_i + c = 0.$$

On applying the theorem of (5, 25) we get easily the following types as affine normal forms.

$$q(x) = 0, \qquad r = 1, \ldots, n \qquad r/2 \leq p \leq r$$

$$q(x) = 1, \qquad r = 1, \ldots, n, \qquad p \leq r \tag{5, 30}$$

$$q(x) = x_{r+1}, \qquad r = 1, \ldots, n-1, \qquad r/2 \leq p \leq r$$

If we replace $q(x)$ by the quadratic form $Q(x)$ of (5, 27) we get the types of quadrics different in the sense of metric geometry. The formula (5, 27) can also be interpreted for projective geometry.

Let $R(x_1, \ldots, x_n)$ and $S(x_1, \ldots, x_n)$ be two quadratic forms, S representing a quadric without real points. Then we transform the coordinates in such a manner that S is transformed to $S = x_1^2 + \ldots + x_n^2$, and R to R'. By any orthogonal transformation S will not be altered, but we can transform R' by an orthogonal transformation to the normal-form (5, 27). Hence we can transform simultaneously

R to $a_1 x_1^2 + \ldots + a_n x_n^2$,

S to $x_1^2 + \ldots + x_n^2$ and the pencil

$\kappa R + \lambda S$ to $(\kappa a_1 + \lambda) x_1^2 + \ldots + (\kappa a_n + \lambda) x_n^2$

The elements of the the pencil have therefore been transformed to the normal-form simultaneously.

§ 6. Resultants

Let {6/1}

$$f(x) = x^n + a_1 x^{n-1} + \ldots + a_n = (x - \alpha_1) \ldots (x - \alpha_n) \tag{6, 1}$$

$$g(x) = x^m + b_1 x^{m-1} + \ldots + b_m = (x - \beta_1) \ldots (x - \beta_m) \tag{6, 2}$$

$$R(f, g) = \prod_i \prod_k (\alpha_i - \beta_k). \tag{6, 3}$$

Then $R(f, g)$ is uniquely defined by the polynomials f and g, and it is said to be the *resultant* of f and g. The necessary and sufficient condition that f and g may have a common root is that the resultant is equal to zero.

From (6, 3) it follows

$$R(f, g) = (-1)^{mn} R(g, f) ; \tag{6, 4}$$

from (6, 2) and (6, 3)

$$R(f, g) = \prod_{\nu=1}^{n} g(\alpha_\nu), \tag{6, 5}$$

and by interchanging f and g and applying (6, 4) we get

$$R(f, g) = (-1)^{mn} \prod_{\mu=1}^{m} f(\beta_\mu). \tag{6, 6}$$

The right side of the equation (6, 5) is a symmetric polynomial in $\alpha_1, \ldots, \alpha_n$ with coefficients $b_0, b_1, \ldots, b_m$, these polynomials having integral coefficients. From Part II (1, ...) it follows that $R(f, g)$ can be represented as a polynomial in the elementary symmetric polynomials of $\alpha_1, \ldots, \alpha_n$ with coefficients $b_0, b_1, \ldots, b_m$, where ... has integral coefficients. As the elementary symmetric polynomials of $\alpha_1, \ldots, \alpha_n$ are equal to ..., the resultant $R(f, g)$ can be represented as a polynomial in $a_1, \ldots, a_n, b_1, \ldots, b_m$ with integral coefficients.

For a reason which will become evident later on this representation will be written as

$$R(f, g) = R(1, a_1, \ldots, a_n; 1, b_1, \ldots, b_m). \tag{6, 7}$$

If in any term $A = a_1^{s_1} a_2^{s_2} \ldots a_n^{s_n} b_1^{t_1} b_2^{t_2} \ldots b_m^{t_m}$ the factors a_ν are represented by $\alpha_1, \ldots, \alpha_n$ and the factors b_μ by $\beta_1, \ldots, \beta_m$, then A becomes an homogeneous polynomial in α_ν and β_μ of degree *

$$t = s_1 + 2s_2 + \ldots + n s_n + t_1 + 2t_2 + \ldots + m t_m \tag{6, 8}$$

which is called the weight of A. From (6, 5) it follows that $R(f, g)$ is homogeneous of degree mn; hence each term of $R(1, a_1, \ldots, a_n; 1, b_1, \ldots, b_m)$ has the weight mn. From (6, 5) we see that one of those terms is equal to b_m^n.

Let S be a polynomial in $a_1, \ldots, a_n, b_1, \ldots, b_m$ with the property that S becomes equal to zero if f and g have a common root. By representing a_ν as symmetric polynomials of $\alpha_1, \ldots, \alpha_n$ and b_μ as symmetric polynomials

* See Part II, p. 39

in $\beta_1, \ldots, \beta_m$ we get

$$S = \Sigma = \Sigma(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m)$$

The right side of this equation is equal to zero if $\alpha_1 = \beta_1$. Hence

$$\Sigma = \Sigma(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m) - \Sigma(\alpha_1, \ldots, \alpha_n, \alpha_1, \ldots, \beta_m)$$

Subtracting the corresponding terms in the right side we see that Σ is divisible by $\alpha_1 - \beta_1$ and in the same manner that Σ is divisible by $\alpha_i - \beta_k$, hence Σ is divisible by $R(f, g)$. [illegible] From the 2nd theorem of Part II (10), it follows therefore that $R(1, a_1, \ldots, a_n; 1, b_1, \ldots, b_m) = S$. $R(1, a_1, \ldots, a_n, 1, b_1, \ldots, b_m)$. Hence S is divisible by the resultant. The weight of every term is therefore not less than $m n$. If each term has the weight $m n$, S differs from the resultant by a factor of weight 0 only and this factor is the coefficient of the term $b_m{}^n$ in S.

To get the resultant as a polynomial in $a_1, \ldots, a_n, b_1, \ldots, b_m$ we have therefore to find out a polynomial S, with the following three properties. [6/2]

1. $S = 0$ if f and g have a common root.
2. Each term of S has the weight $m n$.
3. The term $b_m{}^n$ has the coefficient 1.

A polynomial of this kind can easily be found out by the following consideration.

Let $f(\alpha) = 0 = g(\alpha)$, then the following $n + m$ equations hold

$$\alpha^{m-1} f(\alpha) = \alpha^{n+m-1} + a_1 \alpha^{n+m-2} + \ldots + a_n \alpha^{m-1} = 0$$

$$\alpha^{m-2} f(\alpha) = \qquad \alpha^{n+m-2} + \ldots + a_{n-1} \alpha^{m-1} + a_n \alpha^{m-2} = 0$$

$$\ldots$$

$$f(\alpha) = \qquad \alpha^{n} + a_1 \alpha^{n-1} + \ldots + a_n = 0$$

$$\alpha^{n-1} g(\alpha) = \alpha^{n+m-1} + b_1 \alpha^{n+m-2} + \ldots + b_m \alpha^{n-1} = 0$$

$$\alpha^{n-2} g(\alpha) = \qquad \alpha^{n+m-2} + \ldots + b_{m-1} \alpha^{n-1} + b_m \alpha^{n-2} = 0$$

$$\ldots$$

$$g(\alpha) = \qquad \alpha^{m} + b_1 \alpha^{m-1} + \ldots + b_m = 0$$

We consider this system as a system of $m+n$ linear equations in $x^{m+n-1}, \ldots, x$. It can be satisfied only if its determinant is equal to zero. Hence a necessary condition that f and g may have a common root is

$$S = \begin{vmatrix} 1 & a_1 & \cdots & a_n & & & \\ & 1 & \cdots & a_{n-1} & a_n & & \\ & & \cdots & & & & \\ & & & 1 & a_1 & \cdots & a_n \\ 1 & b_1 \cdots & b_m & & & & \\ & 1 \cdots & b_{m-1} & b_m & & & \\ & & \cdots & & & & \\ & & & 1 & b_1 & \cdots & b_m \end{vmatrix} = 0 \qquad (6,8)$$

The terms of S are of weight mn each; the term b_m^n is the diagonal element and has the coefficient 1. Hence

$$S = R(f, g). \qquad (6,9)$$

[6,5] Let $F(x) = a_0x^n + \ldots + a_n$

$G(x) = b_0x^m + \ldots + b_m$,

where nothing is supposed about the coefficients. We define now

$$R(F, G) = R(a_0, \ldots, a_n, b_0, \ldots, b_m)$$

$$= \begin{vmatrix} a_0 & \cdots & a_n & & \\ & \cdots & & & \\ & & a_0 & \cdots & a_n \\ b_0 & \cdots & b_m & & \\ & \cdots & & & \\ & & b_0 & \cdots & b_m \end{vmatrix} \qquad (6,10)$$

As in (6,8) there are m rows with elements a_i and n rows with elements b_i. If $a_0 = b_0 = 1$, then $F = f$, $G = g$, and we see from (6,9) and (6,10), that the notation $R(F, G)$ conforms to $R(f, g)$. We have to consider three cases.

1. $a_0 \neq 0$, $b_0 \neq 0$.

$$F(x) = a_0\left(x^n + \ldots + \frac{a_n}{a_0}\right) = a_0\varphi(x) = a_0(x - \alpha_1) \ldots (x - \alpha_n)$$

$$G(x) = b_0\left(x^m + \ldots + \frac{b_m}{b_0}\right) = b_0\psi(x) = b_0(x - \beta_1) \ldots (x - \beta_m)$$

From (6,10) it follows that*

$$R(F, G) = a_0^m b_0^n R(\varphi, \psi). \qquad (6, 11)$$

Hence for (6, 3) (6, 5), and (6, 6)

$$R(F, G) = a_0^m b_0^n \prod_i \prod_k (\alpha_i - \beta_k) = [illegible] \qquad (6, 12)$$

Hence in this case R(F, G) = 0 is the necessary and sufficient condition for the existence of a common root of F and G.

2. $a_0 = 0$, $b_0 = 0$. From (6,10) it follows that R(F, G) = 0, independent of the existence of a common root.

3. $a_0 \neq 0$, $b_0 = 0$ (or $a_0 = 0$, $b_0 \neq 0$).

Let $b_1 = \dots = b_m = 0$, then R(F, G) = 0 and every root α' of F satisfies obviously $G(\alpha') = 0$.

Let every $b_{i<r} = 0$ [illegible] $\neq 0$ [illegible] $+ \dots + b_m$ for [illegible]

By setting 0 for b_0 in (6, 10) we get R(F, G) $= a_0$ R(F, G [illegible])

and by mathematical induction [illegible] R(F, G [illegible]) [illegible] R(F, G)

Hence R(F, G) = 0 if and only if F and G have a common root [illegible] if F and G have a common root [illegible] the corresponding proposition holds for [illegible] $b_r \neq 0$. By these considerations we get the following theorem.

Theorem [illegible] If F and G have a common root, then R(F, G) = 0. If R(F, G) = 0, then either $a_0 = b_0 = 0$ or F and G have a common root.

Exercises. 1. Consider the [illegible] (6, 11) and (6, 12) in the case 3.

2. Let $F(x, y) = \sum [illegible]$ and $G(x, y) = \sum b_k [illegible] y^k$.

State the necessary and sufficient condition for R(F, G) $\neq 0$.

Let [illegible] be the factors of the last column of the [illegible] determinant (6, 10), and let

$$u(x) = u_1 x^{m-1} + \dots + u_{m-1} x + u_m, \quad v(x) = v_1 x^{n-1} + \dots + v_{n-1} x + v_n.$$

* See Part I, p. 22

15

When we multiply the $n+m$ equations

$$
\begin{aligned}
x^{m-1}F(x) &= a_0 x^{n+m-1} + \dots + a_n x^{m-1} \\
&\dots\dots\dots\dots\dots\dots \\
F(x) &= \qquad\qquad a_0 x^n + \dots + a_n \\
x^{n-1}G(x) &= b_0 x^{n+m-1} + \dots + b_m x^{n-1} \\
&\dots\dots\dots\dots\dots\dots \\
G(x) &= \qquad\qquad b_0 x^m + \dots + b_m
\end{aligned}
$$

with $u_1, \dots, u_m, v_1, \dots, v_n$ respectively and add, we get

$$u(x)\,F(x) + v(x)\,G(x) = R(F, G). \qquad (6, 13)$$

We can express this formula by the following theorem.

Theorem. Let R be the ring generated by the indefinite x and the coefficients of F and G, then R(F, G) is linearly dependent on F and G with coefficients from R.

Exercise. Prove the theorem of [6/3], without any reference to symmetric functions, by the help of the last theorem. (Special attention should be given to the case when every cofactor is equal to zero.)

If $F(x)$ and $G(x)$ have a common root, the highest common factor (F, G) is a polynomial of positive degree; we can get it by the algorithmus of the *h.c.f.*, hence its coefficients belong to the ring generated by the coefficients of F and G by addition, subtraction and multiplication.

[6/5] Let the coefficients $a_0, \dots, a_n$; $b_0, \dots, b_m$ of $F(x)$ and $G(x)$ be polynomials of $K(y)$, K being an arbitrary field;

$$F(x) = f(x, y), \quad G(x) = g(x, y).$$

The resultant R(F, G) is therefore a polynomial in y, and from (6, 13)

$$R(F, G) = R(y) = u(x, y)\,f(x, y) + v(x, y)\,g(x, y). \qquad (6, 14)$$

We suppose that at least one of the polynomials $a_0 = a_0(y)$, $b_0 = b_0(y)$ is different from the polynomial 0. From [6/3] it follows that $f(x, y)$ and $g(x, y)$ have a common factor depending on x, if and only if $R(y)$ is the polynomial 0. The procedure of getting $R(y)$ and $g(x, y)$ is called *elimination of x*. Let η be a root of $R(y)$; then $R(f(x, \eta), g(x, \eta)) = R(\eta) = 0$. Hence either $a_0(\eta) = b_0(\eta) = 0$, or $f(x, \eta)$, $g(x, \eta)$ have a common root. By this method we can find out the common solutions of two equations

$$f(x, y) = 0, \qquad g(x, y) = 0.$$

CORRECTIONS

Part I. (see the corrections given in Part II.)

Page.	*Line.*	*Read*	*For*
iii (Preface)	10	does	do
13	19	(2/H)	(2)
16	14	depend on	apply to
20	10	10.	(11)
	23	11.	10.

Part II.

Page.	*Line.*	*Read*	*For*
6	10	$e + O$	$e + O$
6	19	M	M'
11	22	$(b' + w)$	$(b' + t)$
	23	$a'w$	$a't$
13	20	are	and also the distributive law are
	22	a non-distributive system	a system
	28	the reader may verify that (2, 8) generates an addition and a multiplication of the classes for which the commutative, associative, and distributive laws hold, and	we should only prove that
18	21	field	field
18	11	e_{n-1}	e_{n+1}
21	12 and 13 : interchange the exponents 1		
23	24	$f(x), \phi(x), \psi(x)$	$f(x), \psi(x), \psi(x)$
26	10	$\frac{a_n}{b_m}$	$\frac{b_m}{a_n}$
	11	$\frac{a_{n_2}}{b_m}$	$\frac{b_m}{a_{n_2}}$
	26	$\phi(x)\,\phi_1(x)$	$\phi(x)\psi_1(x)$
33	6	K	K_1
	20	K(θ)	K(a)
38	9	N	K [three times !]

Page.	*Line.*	*Read*	*For.*
42	18	to K(a)	to K
46	9	a_s	a_t
47	19	$x^{n-1} + \dots + b_n$	$x^{m-1} + \dots + b_m$
	30	$F(a_1, \dots, a_n)$	$F(a_1, \dots, a_m)$
52	17	h	h [error not in all copies]
61	3	(13, 2)	(13, 2
	5	(13, 2)	(13, 1)
	6	$n(n-1)$	$n(n-1)$

Parts III—V. (in this volume)

Page.	*Line.*	*Read*	*For.*
3	19	P′	P [twice]
	20	Q′	Q [twice]
	25	a_{i+j+1}	a_{i+j-1}
8	25	$s-\nu$	$s+1-\nu$
11	5	$>a>$	$<a<$
12	4	$+$	$-$ [between the fractions]
13	27	$\frac{P_{2n}}{Q_{2n}}$	$\frac{P_{2n}}{P_{2n}}$,
14	17	$s_n \mid \lambda$	s_n, λ
16	1	α	a
	15, 19	purely periodic	purely
19	8	[8/4] [on the margin]	
	9	>1	>0
20	12	$\alpha+\beta$	$\alpha_1+\beta_1$
		$s+\alpha$	$s+\alpha_1$ [twice]
22	5	$\sqrt{26}$	26
31	1	Σs_i	s_i
	6	If this sum is convergent, the sum taken for odd indices is divergent, where it follows easily that $Q_{2n-1} \to \infty$. Hence	Hence
	10	interchange "then" with "and"	

Page.	*Line.*	*Read*	*For*
41	17	(1, 3)	(3)
42	21	$(x-a)(x-\bar{a})$	$(x-\bar{a})(x-\bar{a})$
58	16	(1, 30)	(30)
58	5	2·40824	2·40224
59	15	$\left(1+\frac{2b_3^m+b_4^m}{b_1^m}\right)$	$\frac{(1+2b_3^m+b_4^m)}{b_1^m}$
61	14	b_1	b_4
66	15	[6/2]	[6/]
68	22	Δ_{s-t}	Δ_{st}
72	7	(see Part II [1/2])	(see Part II [1/2],
	18	modul M	modul *m*
	30	$((e^1_j))$, where $e^1_1=1$, and e^1_j	(e^1_j), where $a^1_1=1$, and a^1_j
74	7	exists	exist
75	18	e^1_j	e^1_j
76	21	exists	exist
79	26	[omit =]	
80	10	[place λ from the 2nd to the 1st column in the 1st line of (2, 14)]	
	11	$A^{(r)}$	$A^{(n)}$
		$\chi_{A^{(r)}}$	$\chi_{A^{(k)}}$
82	14	[replace the index *m* by *q*]	
85	9	$W_{1,r}$	$W_{1,r-1}$
101	17	exists	exist
105	23	$(x)^n A(y)$	$A(y)$